THE UNIVERSITY *of* EDINBURGH

The cellulose degradation system of *Cytophaga hutchinsonii*

Chao-Kuo Liu

PhD

The University of Edinburgh

2012

DECLARATION

I declare that this thesis was composed by myself. The work presented in this thesis is my original work and has not been submitted for any other degree or professional qualification except as specified.

Chao-Kuo Liu

Acknowledgements

I would like to thank Dr. Chris French and Dr. Andrew Free for their great guidance and supervision. I would also like to thank Prof. Stephen Fry and Prof. Neil Bruce for being my examiners with their helpful suggestions and advices. Thanks to Dr. Bruce Ward, Dr. Maurice Gallagher, all the members in Chris' Lab and all the people in the 7th and 8th floor of Darwin Building for their helps and lovely supports.

Special thanks to Mr. Tim Gregson for the HPLC analysis, and the financial support from the School of Biological Sciences, University of Edinburgh.

I would also like to give my full appreciation to Dr. Hsiu-Ping Pearl Lin who gives me a chance to enter the world of molecular biology.

Finally, I would like to dedicate this thesis to all my family members for their full supports and understandings during the past years.

Abstract

In this project, *Cytophaga hutchinsonii*, an aerobic gliding bacterium with cellulose-degrading ability, was studied, since its cellulase system was unknown and might be very different from those of other cellulose-degrading species. Only ß-1,4-endoglucanases and non-specific ß-glucosidases were found in the *C. hutchinsonii* genome sequence, whereas specific exoglucanases were apparently absent. Almost all putative cellulases were composed of catalytic domains only, without carbohydrate-binding modules. Samples from *C. hutchinsonii* cultures were analyzed by using TLC and colorimetric assays. Glucose was detected in the cellobiose grown culture, but not in cellulose-grown cultures, suggesting that cellobiose is hydrolyzed extracellularly rather than being directly assimilated, and that cellulose may not be degraded via cellobiose. Also, cellobiose-based cultures caused greater acidification of the medium than glucose or cellulose grown cultures. Nine putative cellulases were expressed in four bacterial strains. In some cases, expression was toxic to host cells. The crude lysates were tested for endoglucanase, specific exoglucanase or non-specific ß-glucosidase activity. CHU_1280 and CHU_1842 showed apparent endoglucanase activity when expressed in *Citrobacter freundii*. Four putative GH family 3 ß-glucosidases with similar conserved domains were expressed in *Escherichia coli* JM109 and *E. coli* BL21(DE3)pLysS. One of these, CHU_2268, was found to possess MUC-degrading ability. This suggests that CHU_2268 may be the 'missing' exoglucanase in *C. hutchinsonii*. Another two ß-glucosidases, CHU_2273 and CHU_3784, possessed only MUG-degrading activity.

Table of Contents

List of Tables

List of Figures

Abbreviations

BM	Binding Module
CAZy	Carbohydrate-Active enZYmes (Database)
CBD	Cellulose-binding Domain
CBM	Carbohydrate-binding Module
CD	Catalytic Domain
CMC	Carboxymethyl Cellulose
EPS	Extracellular Polysaccharide(s)
GH	Glycoside Hydrolase
GT	Glycosyltransferase
HPLC	High-performance Liquid Chromatography
KEGG	Kyoto Encyclopedia of Genes and Genomes
LB	Luria-Bertani (Medium)
MFS	Major Facilitator Superfamily
4-MU	4-Methylumbelliferone
MUC	4-Methylumbelliferyl-ß-D-cellobioside
MUG	4-Methylumbelliferyl-ß-D-glucopyranoside
NCBI	National Center for Biotechnology Information
NC-IUBMB	Nomenclature Committee of the International Union of Biochemistry and Molecular Biology
OD	Optical density
PBS	Phosphate Buffer Saline
SE	Standard Error
SEM	Scanning Electron Microscope
SUS	Starch Utilization System
TEM	Transmission Electron Microscope
TLC	Thin Layer Chomatography

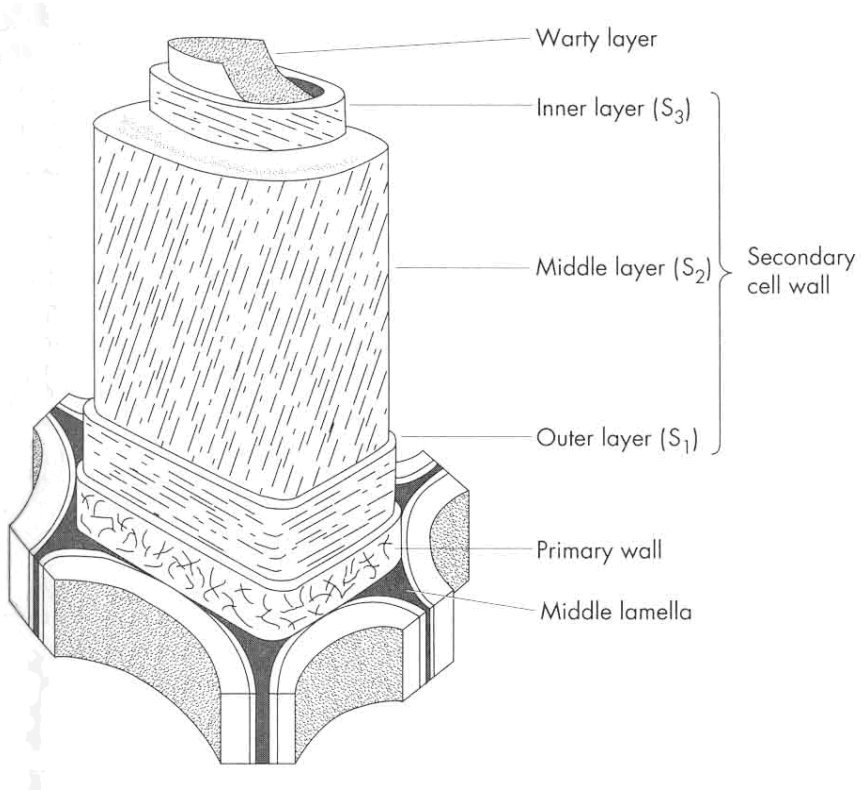
Chapter 1 Introduction

Cellulose degradation by microorganisms is a fascinating research topic, with increasing importance in industrial biotechnology. During the last century, scientists have isolated cellulose-hydrolyzing microorganisms and sought to discover which enzymes (cellulases) might be involved in cellulose degradation and what kind of enzymatic systems exist in these microorganisms. Because of the advances in molecular biology and biotechnology, scientists can now analyze cellulase structure in detail and discover novel cellulose-degrading mechanisms. Recently, scientists have even modified *Escherichia coli*, a non-cellulolytic microorganism, to degrade cellulose and produce monoterpene pinene (an immediate chemical precursor to a potential jet fuel), n-butanol and amorphadiene directly [Bokinsky et al. 2011, Mckee et al. 2012]. In the research described in this thesis, an early-discovered cellulose-degrading microorganism, *Cytophaga hutchinsonii*, was analyzed and some of its cellulases were examined.

1.1 Abundance and structure of cellulose

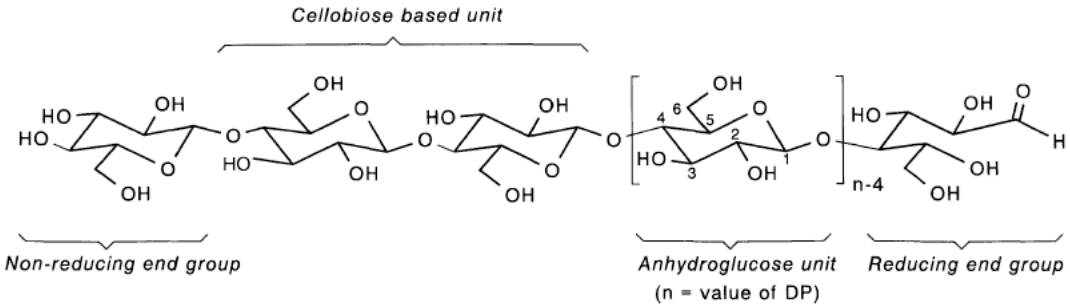
Cellulose is the most abundant component of plant biomass [Clarke 1997, Glazer & Nikaido 1998]. It is the most basic cell wall material in the plant and it constructs the framework structure of cell walls in the form of crystalline microfibrils [Fujita & Harada, 2001]. The primary plant cell wall is the first-formed wall layer. Its cellulosic microfibrillar skeleton is deposited while the cell is (at least potentially) still expanding [Fry, 2004] and its thickness in mature cells depends on the cell type

[Harris & Stone 2008]. When cell growth has ceased, hydroxycinnamic acid-mediated cross-linking may occur between the cell wall constituents, and then the secondary wall is formed [Boudet, 2003]. The secondary cell wall consists of three layers, known as S1, S2, and S3 (Figure 1-1). S1 is the outer layer of the secondary cell wall, and is about 0.2 – 0.3 μm in thickness. S1 and S2, the middle layer and the thickest layer of the cell wall (usually 1 – 5 μm), contain the most abundant cellulose [Klemm et al. 1998]. All three S layers consist of cellulose microfibrils, cross-linked with a matrix of hemicellulose and lignin. Usually, cellulose forms 45 – 55% dry weight of softwoods and hardwoods [Glazer & Nikaido 1998].



(Source: Glazer & Nikaido 1998)

Figure 1-1. Structure of cell wall layers.



(Source: Klemm et al. 1998)

Figure 1-2. Basic cellulose structure.

Cellulose is a linear polymer of ß-(1,4)-linked anhydrous glucose residues, which may extend to 15,000 residues in length [Clarke 1997]. The basic repeating unit of cellulose is cellobiose, which consists of two ß-D-glucopyranoses with 1,4-glycosidic linkage so that one monosaccharide is rotated 180 degrees to its neighbor (Figure 1-2). Two hydrogen bonds from one glucose residue to its adjacent glucose residue – one bond between the C6 hydroxyl and the C2 hydroxyl and one bond between the C5 oxygen and C3 hydroxyl – stabilize the glycosidic bond and make the structure stiff. There are also hydrogen bonds between cellulose chains to form a sheet-like structure [Henriksson & Lennholm 2009]. Therefore, intramolecular hydrogen bonds and van der Waal's force strengthen the chain formation of cellulose and intermolecular hydrogen bonds cross-link different chains of cellulose to form microfibrils [Teeri 1997, Zhang & Lynd 2004]. Cellulose microfibrils with diameters as small as 2 nm are found in the primary walls of some plants, whereas microfibrils 10 times wider occur in some seaweed cell walls and tunicates [Bayer et al. 1998a]. Most of the 'amorphous phase' of cellulose corresponds to chains that are located at the microfibril surface, whereas crystalline components occupy its core [Bayer et al. 1998a, French & Johnson 2007]. The most 'realistic' model of the cellulose amorphous region was elaborated in the 1950's by Hearle, supported by wide-angle

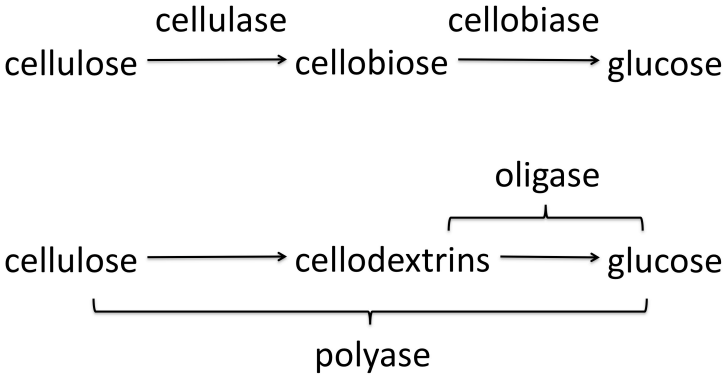
X-ray scattering and ^{13}C high resolution cross-polarization magic angle spinning solid state NMR spectrum tests. In this model, there is a non-uniform fringe fibrillar structure with ordered and disordered regions differing in size and perfection of the crystallites [Klemm et al. 1998]. The degree of crystallinity, degree of polymerization (DP), and width of cellulose microfibrils depend on its source, age and pretreatment [Clarke 1997]. For example, in cotton fiber, DP values are about 800 to 10000, where in wood pulp, DP values are typically about 300 to 1700 [Klemm et al. 2005].

1.2 Microbial cellulose degradation

1.2.1 History of cellulase research

The concept of cellulase originated in the late 19th century. Cellulose–degrading enzyme activity was initially named 'cytase' or 'cytohydrolyst' [Newcombe 1899]. In 1912, Prof. Hans Pringsheim first used the word 'cellulase' in a report published in the German journal 'Hoppe-Seyler´s Zeitschrift für Physiologische Chemie' [Clarke 1997, Sheehan & Himmel 1999]. Large-scale research on cellulase was not started until the 1940s. During World War II, a wide-ranging project was initiated by the U.S government to develop pesticides or reagents against military equipment rot due to the fact that U.S. Marine soldiers' clothes, tents, and even belts were disintegrating in the South Pacific theatre. This project coincidentally improved research on cellulose hydrolysis and had a long-term impact on the history of cellulase science. During this project, one fungus was isolated from a rotting cartridge belt in New Guinea by the U.S. Army Natick Development Center. This was found to possess outstanding cellulose degradation

ability, and was named *Trichoderma viride* [Augustine 1976]. The wild type strain QM6a was later re-identified and renamed by Dr. Emory G. Simmons as *Trichoderma reesei*. *T. reesei* has become one of the most important and best-studied cellulose degrading species in the world [Lizon & Samuels 1997, Montenecourt & Eveleigh 1977].



(Re-drawn from Stone 1958)

Figure 1-3. Early models of cellulose degradation systems.

Stanier (1942) proposed that two enzyme systems, cellulase and cellobiase, were responsible for cellulose decomposition and this concept was generally accepted by scientists at that time. Reese et al. (1950) reported that cellulose degradation was accomplished by two enzymes, C_1 and Cx, rather than cellulase or cellobiase. Stone (1958) suggested that cellulose hydrolysis also required an 'oligase' which might degrade cellodextrins to cellobiose, a suggestion which arose from studies of *Aspergillus* spp. (Figure 1-3). However, more and more evidence suggested that cellulose degradation required multiple enzymes working together. Mandels and Reese (1959) considered that cellulase was an 'extracellular enzyme complex' produced by all cellulolytic fungi. King and Vessal (1969) suggested that the 'cellulase complex' consisted of several enzymes: C1, ß-glucosidase and ß-

(1→4) glucanase [equivalent to Cx, and it had two sub-groups: exo-ß-(1→4) glucanase and endo-ß-(1→4) glucanase]. Currently, almost all scientists agree that cellulase consists of three groups of enzymes: ß-1,4-endoglucanase (EC 3.2.1.4), exoglucanases (EC 3.2.1.74, EC 3.2.1.91 and EC 3.2.1.176) and ß-glucosidase (3.2.1.21) [Clarke 1997, Lynd et al. 2002, Philippidis 1994].

1.2.2 Functions of microbial cellulases

As described in the preceding section, cellulase can be divided into three groups: ß-1,4-endoglucanase (EC 3.2.1.4), exoglucanase (EC 3.2.1.74, EC 3.2.1.91 and EC 3.2.1.176) and ß-glucosidase (EC 3.2.1.21). These cellulases are widespread in archaea, bacteria, fungi, yeasts, plants or animal digestive systems. In this section, only microbial cellulases are introduced.

1.2.2.1 ß-1,4-Endoglucanase

According to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB), ß-1,4-endoglucanase (EC 3.2.1.4) is a cellulase catalysing 'Endohydrolysis of (1→4)-ß-D-glucosidic linkages in cellulose, lichenin and cereal ß-D-glucans'. ß-1,4-Endoglucanase is a cellulase that degrades cellulose at internal amorphous sites in the cellulose polysaccharide chain [Lynd et al.2002]. Warren (1996) suggests that the exo-acting enzymes (exoglucanase) remove units of one or more sugars from the ends of polysaccharide chains, where the endo-acting enzymes (ß-1,4-endoglucanase) hydrolyze random bonds within the chains, thereby producing more ends for the exo-enzymes to act on. Therefore, ß-1,4-endoglucanase is essential for macro-structural cellulose

degradation. Almost all microbial ß-1,4-endoglucanases are extracellular; however, Yamane et al. (1970) and Morana et al. (2008) have reported membrane-associated ß-1,4-endoglucanases from *Pseudomonas fluorescens* var. *cellulosa* and *Alicyclobacillus acidocaldarius*. ß-1,4-Endoglucanase in Cel9R from *Clostridium thermocellum* was reported to release cellotetraose (main product), cellotriose, cellobiose and glucose after hydrolysis of amorphous cellulose [Zverlov et al. 2005]. Although ß-1,4-endoglucanase is well known for its ability to attack the amorphous part of cellulose fibrils [Lynd et al. 2002], some ß-1,4-endoglucanases not only degrade CMC or acid-swollen cellulose, but also avicel (microcrystalline cellulose) [Gilad et al. 2003]. ß-1,4-Endoglucanases have been reported to be inhibited by cellobiose [Gruno et al. 2004, Stoppok et al. 1982], Cu^{2+} [Kim 1995], Hg^{2+} [Fugino et al. 1989, Gruno et al. 2004, Liu & Xia 2006] or other chemicals. The activity of ß-1,4-endoglucanase can be tested by measuring the weight loss of insoluble substrates, changes in turbidity of cellulose suspensions, increase in reducing end groups, decrease in the viscosity of cellulose derivatives, colorimetric determination and measurement of clearance zones in cellulose agar [Eriksson & Pettersson 1988]. Different kinds of substrates are tested to show ß-1,4-endoglucanse activity. Carboxymethyl cellulose (CMC) [Fujii et al. 2009, Sul et al. 2004], acid swollen cellulose [Li & Wilson 2008, Sheweita et al. 1996], filter paper [Kim 1995] and dyed cellulose [Moore et al. 1979] are common substrates used for assays of ß-1,4-endoglucanase activity.

1.2.2.2 Cellobiohydrolase (EC 3.2.1.91) and cellobiohydrolase CelS (EC 3.2.1.176)

Cellobiohydrolase (EC 3.2.1.91) hydrolyses '(1→4)-ß-D-Glucosidic linkages in cellulose and cellotetraose, releasing cellobiose from the non-reducing ends of the chains' [NC-IUBMB]. Teeri (1997) reports that the cellobiohydrolases Cel6A and Cel7A from *T. reesei* are processive and can peel off single surface chains of cellulose. Cellobiohydrolase CelS (EC 3.2.1.176) is a newly created hydrolase group initiated on 30th September 2011 [NC-IUBMB]. This group was created to include cellobiohydrolases which releases cellobiose from the reducing ends of the substrate. Cellobiohydrolase CelS consists mostly of GH (glycoside hydrolase) family 48 cellulases, and the prototypical CelS itself comes from *Clostridium thermocellum* [NC-IUBMB]. CelS is the most abundant subunit of the cellulosome formed by *C. thermocellum* [Wang et al. 1993].

1.2.2.3 Glucan 1,4-ß-glucosidase (EC 3.2.1.74)

Glucan 1,4-ß-glucosidase (EC 3.2.1.74) catalyses 'Hydrolysis of (1→4)-linkages in (1→4)-ß-D-glucans, to remove successive glucose units' and the definition also states that 'Cellobiose is hydrolyzed, but very slowly' [NC-IUBMB]. Glucan 1,4-ß-glucosidase is also referred to as 'cellodextrinase' [Lynd et al. 2002]. It is sometimes reported to show dual ability and is easily confused with ß-glucosidase (EC 3.2.1.21) [Hrmova et al. 1998, Liu et al. 2009]. This type of cellulase hydrolyzes cellodextrins from the non-reducing end [Hrmova et al. 1998, Liu et al. 2009, Qi et al. 2008], and catalytic efficiency is higher when it degrades longer cello-oligosaccharides [Hrmova et al. 1998, Liu et al. 2009]. Glucose is the main

hydrolysis product rather than cellobiose as in the reactions of cellobiohydrolase and cellobiohydrolase CelS.

1.2.2.4 ß-Glucosidase (EC 3.2.1.21)

ß-glucosidase (EC 3.2.1.21) is a hydrolase catalyzing 'Hydrolysis of terminal, non-reducing ß-D-glucosyl residues with release of ß-D-glucose' and included in the definition is the statement that 'Some examples also hydrolyze one or more of the following: ß-D-galactosides, α-L-arabinosides, ß-D-xylosides and ß-D-fucosides' [NC-IUBMB]. Almost all studies indicate that ß-glucosidase can degrade cellobiose, hence it plays a crucial role in cellulose degradation as the ultimate enzymatic step in the biological conversion of cellulose into glucose [Faure et al. 1999, Tenkanen et al. 2003]. Furthermore, as compared with glucan 1,4-ß-glucosidase (EC 3.2.1.74) in the preceding section, both cellulases release glucose; however, whereas ß-glucosidase can hydrolyze cellobiose, glucan 1,4-ß-glucosidase does not act on cellobiose efficiently. ß-Glucosidase is found not only in cellulolytic microorganisms, but also non-cellulolytic microorganisms, such as *E. coli* [Yang et al. 1996]. ß-Glucosidase can be localized in almost every spatial compartment of the cell. ß-Glucosidases can be extracellular [Chen et al. 1994, Yang et al. 2008], membrane-bound [Hayase et al. 2008], periplasmic [Yang et al. 1996], intracellular [González-Pombo et al. 2008] or bound to scaffolding machinery in some anaerobic bacteria [Steenbakkers et al. 2003]. Bhatia et al. (2002) suggest that microbial ß-glucosidase can be sub-divided into three categories: aryl-ß-glucosidase, true cellobiase and broad substrate specificity enzymes. Transglycosylation can also be catalyzed by ß-glucosidase under certain conditions. A ß-glucosidase from the fungus *Piromyces* sp. E2 was

cloned and expressed in *Pichia pastoris*, and transglycosylation was observed, in which cellohexaose and celloheptaose were formed during the degradation test of cellopentaose [Harhangi et al. 2002]. Transglycosylation by ß-glucosidase is also reported in the thermophilic eubacteria *Thermotoga neapolitana* [Turner et al. 2007b], where alkyl-glucosides were synthesized by a GH family 3 ß-glucosidase in an environment containing a high concentration of alcohols and relatively little water.

1.2.2.5 Cleavage of cellulose by oxidoreductases

Cellulases in GH family 61 are widely expressed in fungi [Langston et al. 2011, Phillips et al. 2011], but these cellulases show very weak cellulase activities [Koseki et al. 2008, Cantarel et al. 2009]. The purified GH 61 protein Cel61A (from *T. reesei*) shows much lower activity than Cel7B (cellobiohydrolase) on many polysaccharide substrates assays [Karlsson et al. 2001]. However, recent research articles indicate that the GH family 61 cellulases efficiently catalyze degradation of lignocellulosic biomass by oxidative reactions, rather than hydrolysis, with supplemental metal ions, such as copper, nickel and manganese [Phillips et al. 2011, Quinlan et al. 2011, Westereng et al. 2011]. Neither the carbohydrate binding pocket nor catalytic center of the types normally seen in GH families are identified in the structures of GH 61 enzymes [Harris et al. 2010, Karkehabadi et al. 2008]. Reports also indicate that the oxidative cleavage of polysaccharides by GH 61 enzymes is actually a co-reaction with cellobiose dehydrogenases [Langston et al. 2011, Longoni et al. 2012, Phillips et al. 2011, Sygmund et al. 2012]. In addition to GH 61 cellulases, some members of the CBM 33 (carbohydrate binding-module) family,

such as CBP21 (chitin-binding protein) in the Gram-negative soil bacterium *Serratia marcescens*, also degrade crystalline polysaccharides by oxidation [Forsberg et al. 2011, Vaaje-Kolstad et al. 2010].

1.2.2.6 Functional overlap between cellulases

Although cellulases can be divided into groups as described above, overlap in degradation activity between these groups is reported. For example, Han et al. (1995) reported that a cellulase cloned from a *Bacillus* strain has both ß-1,4-endoglucanase and exoglucanase activities, leaving cellobiose as the end product. Tomme et al. (1996) report that CenC from *Cellulomonas fimi* is able to release mainly cellobiose from soluble cellodextrins and insoluble cellulose. Also, their analysis of CMC hydrolysis suggests that CenC is a semiprocessive enzyme with both ß-1,4-endoglucanase and exoglucanase activities. Cellulase E4 from *Thermomonospora fusca* also showed both endo- and exo-cellulase activities as indicated by enzymatic assays and structural analyses [Sakon et al. 1997, Barr et al. 1996, Irwin et al. 1993].

1.2.3 Microbial cellulase structures and functional diversity

In 1984, Bhikhabhai and Pettersson found that amino acid sequences of some cellulases in *T. reesei* were similar to each other [Henrissat 1994]. In 1986, Tilbeurgh et al. observed that a small polypeptide from *T. reesei* showed cellulose-binding activity [Boraston et al. 2007]. Similar results were observed in CBHII (now Cel6A) from *T. reesei* and a ß-1,4-endoglucanase from *C. fimi* [Gilkes et al. 1988, Gilkes et al. 1989, Henrissat 1994]. Francisco et al. (1993) fused a signal peptide of a major

lipoprotein and a transmembrane domain of OmpA with Cex from *C. fimi* and expressed it on the surface of *E. coli*. They observed that the Cex binding domain (BD) expressed in *E. coli* bound tightly to cellulosic materials. Other studies reported similar results, and it was proposed that a typical cellulase included two domains: a catalytic domain (CD) and a BD, with a linker peptide connecting these two regions (Figure 1-5) [Gilkes et al. 1991, Linder & Teeri 1997, Ståhlberg et al. 1991, Wilson 2009]. However, the structure of cellulases discovered in anaerobic bacteria is more complicated. Anaerobic bacteria, especially in the genus *Clostridium*, have been found to possess a cellulose-degrading complex (cellulosome) in which the CD and BD may be separate. Also, cellulases in plants and termites are commonly found to contain only the CD [Levy et al. 2002, Tokuda et al. 1999].

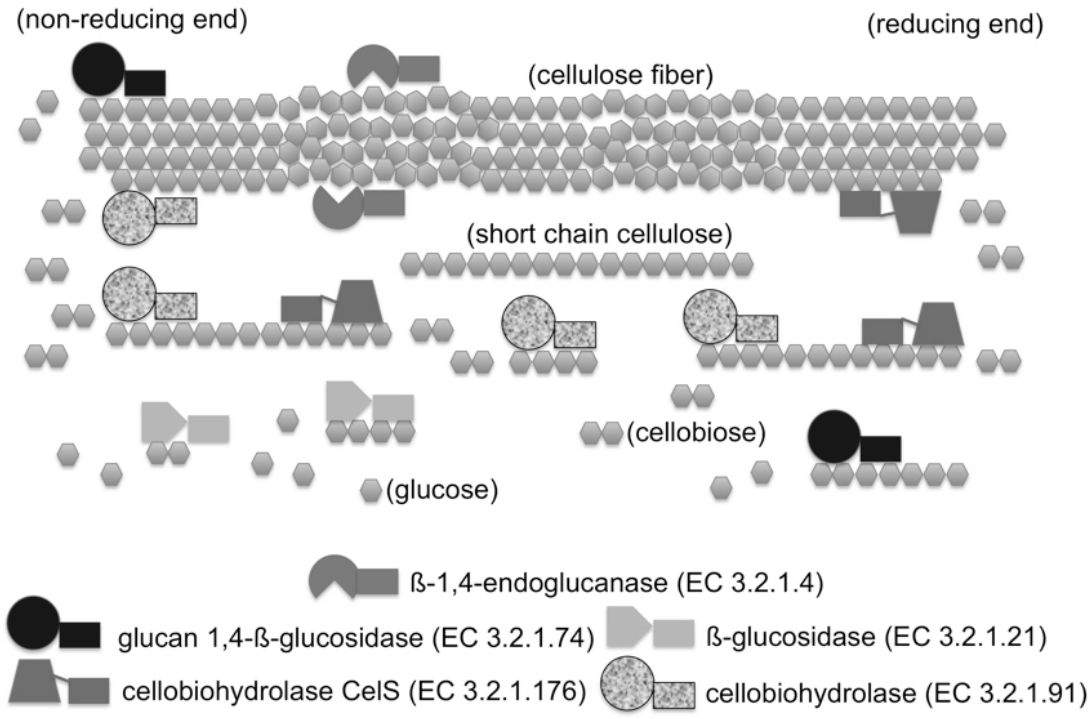
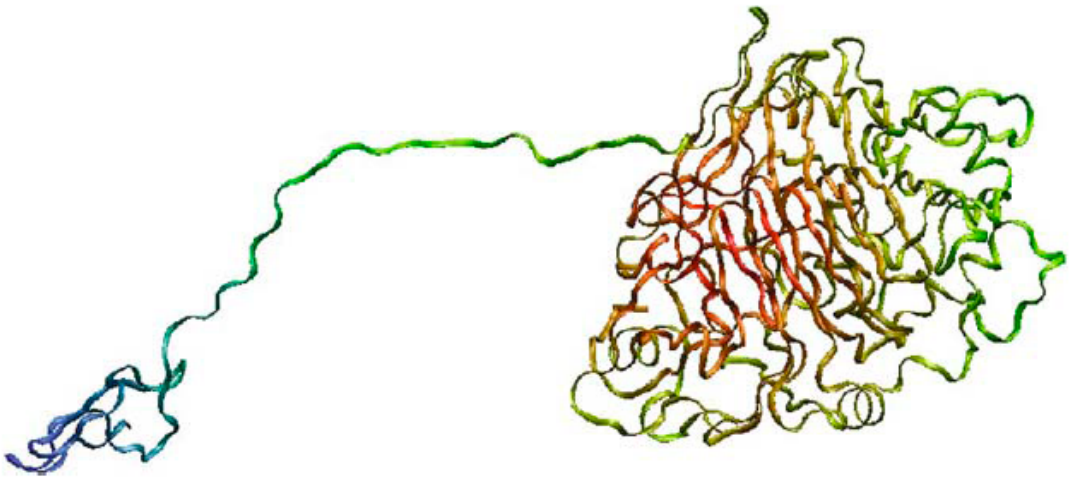


Figure 1-4. Mechanism of non-complexed cellulase system. ß-1,4-Endoglucanase hydrolyzes cellulose fibers from the amorphous regions. The other four cellulases belong to the exo-type hydrolases. Cellobiohydrolase CelS releases glucose from the reducing end.



(Source: Xu et al. 2007)

Figure 1-5. The proposed structure of *T. reesei* CBH I. The structure shows the cellulose-binding module (folded peptide chain on the left, indicated in purple and light blue), a 26-amino acid linker peptide (the light blue and green), and the catalytic domain (the red, orange, yellow and green).

1.2.3.1 Catalytic domains of cellulases

CD and BD in cellulases are now classified into families, according to amino acid sequence similarities [Bayer et al. 1998a]. The CDs of cellulases are included in the GH super-family. There are now 127 GH families. ß-1,4-Endoglucanases (EC 3.2.1.4) can be found in 17 GH families, whereas exoglucanases (EC 3.2.1.74, EC 3.2.1.91, EC 3.2.1.176) and ß-glucosidases (EC 3.2.1.21) can be found in 7 and 6 GH families respectively as shown in Appendix IV. Almost all identified ß-glucosidases are grouped in GH families 1 and 3. There are only five identified microorganisms that produce ß-glucosidases in GH families 9, 30 and 116. The ß-glucosidase of the oomycete *Phytophthora infestans* found in GH family 30 is a bi-functional enzyme which has 4-methylumbelliferyl-ß-D-glucopyranoside and 4-methylumbelliferyl-ß-D-xylopyranoside activity [Brunner et al. 2002]. Studies of CD also show a diversity

of degradation activity. Han et al. (1995) suggest that a cellulase cloned from a *Bacillus* strain has both endoglucanase and exoglucanase activity leaving cellobiose as the end product. A similar result was reported by Li et al. (2007) who suggested that Cel9A in *T. fusca* was a special processive endoglucanase and possessed both endo- and exo-cellulase activities (Figure 1-6).

The reactions of CD are divided in two types: retaining and inverting. Retaining cellulases use a double-displacement mechanism to catalyze hydrolysis with retention of configuration at the anomeric center whereas inverting cellulases use a single-displacement mechanism leading to inversion of configuration at the anomeric center [Clarke 1997].

1.2.3.2 Binding domains of cellulases

The name 'cellulose-binding domain' (CBD) has now been replaced by 'carbohydrate-binding module' (CBM) since binding domains are also found in other polysaccharide hydrolases [Boraston et al. 2007]. The classification of CBM is expanding from 43 families [Boraston et al. 2007] to 55 families [Guillén et al. 2010] and now 64 families [Cantarel et al. 2009]. CBMs are classified into 7 folding families, and structural analysis categorizes CBMs in three groups: (i) 'surface binding' (type A), (ii) 'glycan chain binding' (type B), and (iii) 'small sugar binding' (type C) [Shoseyov et al. 2006]. CBMs are also suggested to possess additional functions, such as non-hydrolytic substrate disruption [Levy et al. 2002] and alteration of interfacial properties of fibers [Pala et al. 2003, Shoseyov et al. 2006].

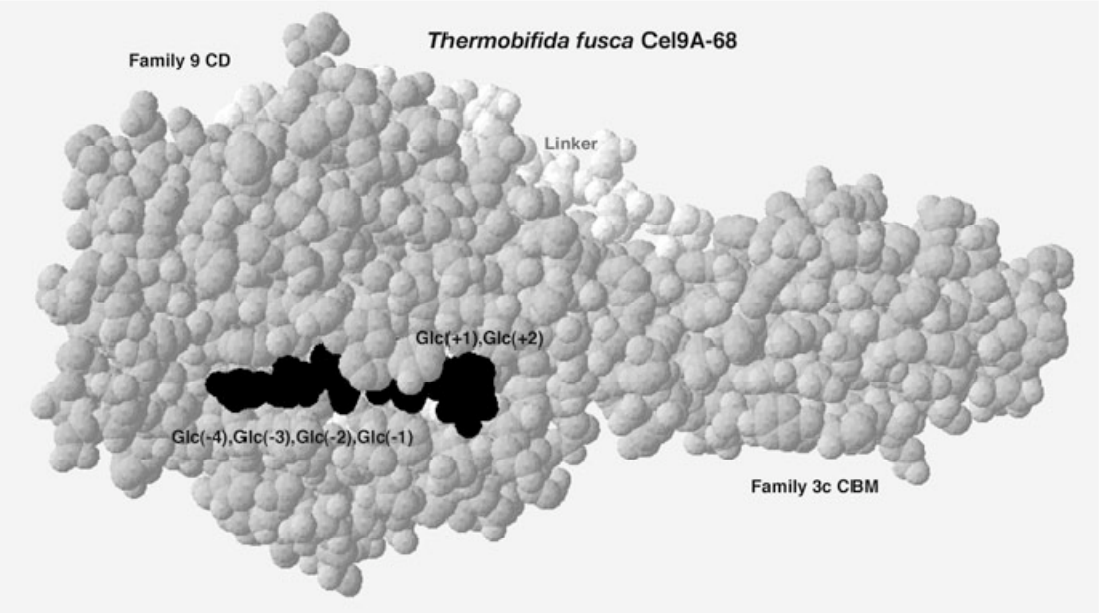
Most studies support the proposal that the CBM is important for cellulose degradation [Hervé et al. 2010]. Hamada et al. (2001) report that the isolated CD of

exoglucanase (Ex1) from *Irpex lacteus* shows lower activity than the original Ex1 and suggest that the CBM is important to enhance hydrolytic activity and binding ability to insoluble cellulose. Denman et al (1996) characterized a cellulase from the rumen fungus *Neocallimastix patriciarum* and reported that the deletion of the CBD results in a marked decrease in the cellulose-binding ability and activity toward crystalline cellulose. However, Nakazawa et al. (2008) cloned three ß-1,4-endoglucanase CD (EGI, II, III) of *T. reesei* and expressed these in several *E. coli* strains. The CMC-Congo Red assay showed that all three cloned domains retain activity against CMC even without the CBM.

1.2.4 Microbial cellulase hydrolysis systems

Although all cellulases cleave a single type of bond in a chemically simplistic substrate, the extensive intermolecular bonding pattern of cellulose generates a fascinating crystalline substrate that is particularly resistant to microbial degradation [Bayer et al. 1998a]. Two different types of cellulase system have been characterized. The 'non-complexed cellulase system' refers to cellulase systems in which cellulases are secreted freely as individual enzymes. This type of system has been found in fungi, actinomycetes and bacteria growing under aerobic conditions. Microbial species like *T. reesei* and *C. fimi* are predicted to use this kind of degradation system. For non-complexed degradation systems, different cellulases target different sites and attack cellulose microfibrils synergistically. ß-1,4-Endoglucanase first attacks the amorphous sites of the cellulose chain, then hydrolyzes and breaks down cellulose to generate cello-oligosaccharides that are further hydrolyzed by exocellulases (Figure 1-4). The cellobiohydrolase and

cellobiohydrolase CelS, from the non-reducing ends and reducing ends, respectively, degrade crystalline cellulose and cello-oligosaccharides to release cellobiose. Glucan 1,4-ß-glucosidase degrades ß-glucans from the non-reducing ends and releases glucose. ß–glucosidase hydrolyzes cellobiose, which is released by exoglucanase, into glucose (Figure 1-4) [Chung et al. 1997, Lynd et al. 2002, Zhang et al. 2006].



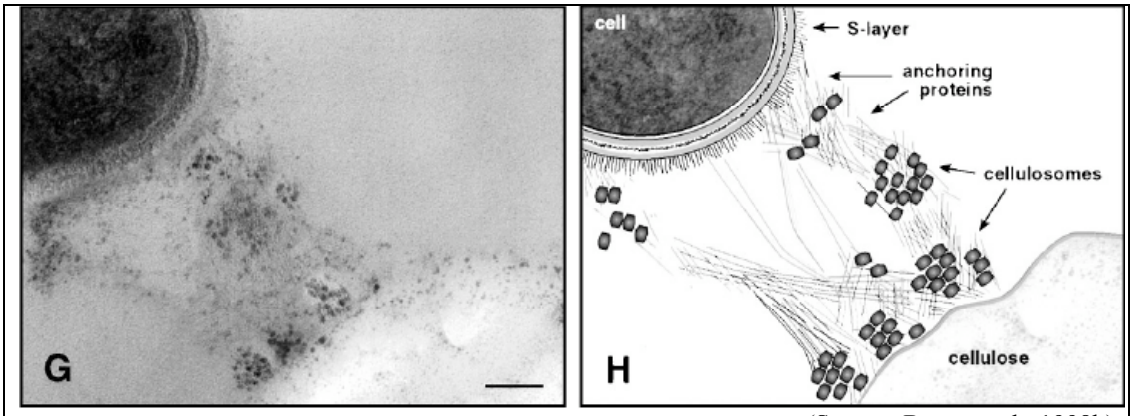
(Source: Wilson 2008)

Figure 1-6. 3D structure prediction of *T. fusca* Cel9A-68. Cel9A-68 possesses a Family 9 CD (catalytic domain), a linker and a Family 3c CBM (carbohydrate-binding module). Its CD forms a tunnel-like active site which binds a cellohexaose molecule (shown in black).

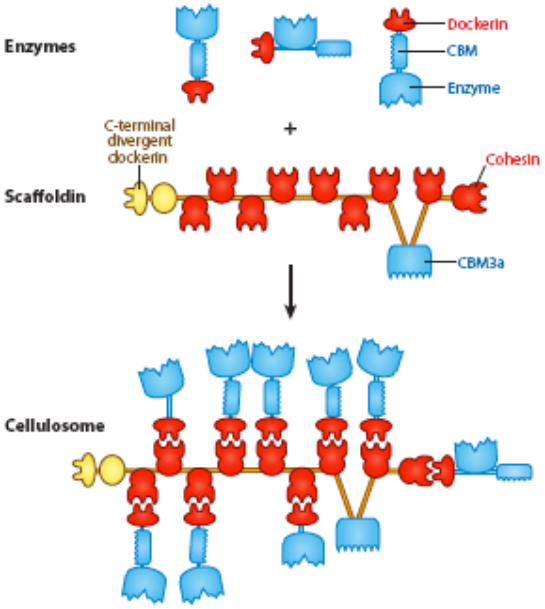
Anaerobic cellulose-degrading bacteria have been found in soils, rumens, termite guts, and artificial environments, such as compost and sewage treatment systems [Doi 2008]. In anaerobic environments, another type of cellulase degradation system occurs, in which cellulases are secreted in a complex called the cellulosome, also known as the complexed cellulase system [Lynd et al. 2002]. Cellulosomes are formed by a series of immobilized cellulases attached to the cell surface. Most cellulosomes are found in anaerobic bacteria, either thermophilic or mesophilic; a few have been found in fungi [Steenbakkers et al. 2003]. The cellulase

complex of anaerobic fungi is less well characterized than the cellulosome in bacteria [Fontes & Gilbert 2010]. The cellulosome is considered to be an extremely efficient complex for cellulose degradation because of spatial co-localization of cellulase activities. However, a recent review [Bayer et al. 2004, Fontes & Gilbert 2010] indicates that cellulosomes are also capable of hydrolyzing heteroxylans, as xylanase has been found associated with cellulosomes in *Clostridium cellulolyticum* and *Clostridium thermocellum* [Mohand-Oussaid et al. 1999, Morag et al. 1990]. Other hemicellulases, such as mannanase and arabinofuranosidase, also occur. The diversity of bacterial cellulosomes has become clear through studies of the family Clostridiaceae, cloning and sequencing of the multiple scaffoldins from *Acetivibrio cellulolyticus* and *Bacteroides cellulosolvens*, and the genome sequencing of *Ruminococcus flavefaciens* [Fontes & Gilbert 2010)].

The cellulosome is comprised from a series of hydrolytic CDs, CBMs, cohesins and dockerins associated by a scaffoldin backbone to bridge cellulases and the cell surface (Figure 1-7). Cellulases (usually CD plus BD) associated with the cellulosome possess a complementary domain called dockerin. Dockerins are usually present in a single copy at the C-terminus of cellulosomal enzymes. The dockerin-appended cellulase binds to cohesin, which is connected to the non-catalytic scaffoldin. An individual CBM sometimes connects to the scaffoldin. One of this kind of CBM belongs to CBM family 3 in *C. thermocellum*, contributes to cellulosome contact directly to the substrate, which improves degradation efficiency, and also leads to cell adhesion. More scaffoldin structures are sometimes found between the cell surface and the cellulase complex [Bayer et al. 2004, Schwarz 2001].



(Source: Bayer et al., 1998b)



(Source: Fontes & Gilbert, 2010)

Figure 1-7. Cellulosomes at the surface of *Clostridium thermocellum* and the mechanisms of cellulosome assembly. The upper image 'G' shows the ultrastructure of the *C. thermocellum* cell surface visualized by TEM. A schematic interpretation is presented as image 'H'. Scale bar in image 'G' represents 100 nm. In the right-hand figure, the modular cellulases and hemicellulases produced by anaerobic microbes contain a dockerin appended to catalytic (enzyme) and noncatalytic CBMs. Dockerins bind the cohesins (red) of a noncatalytic scaffoldin, providing a mechanism for cellulosome assembly. In general, scaffoldins also contain a cellulose-specific family 3 CBM (CBM3a) and a C-terminal divergent dockerin that target the cellulosome to the plant cell wall and the bacterial cell envelope, respectively. The linkers joining the modules in the scaffoldin and catalytic subunits are shown as orange and blue lines, respectively. [Fontes & Gilbert 2010]

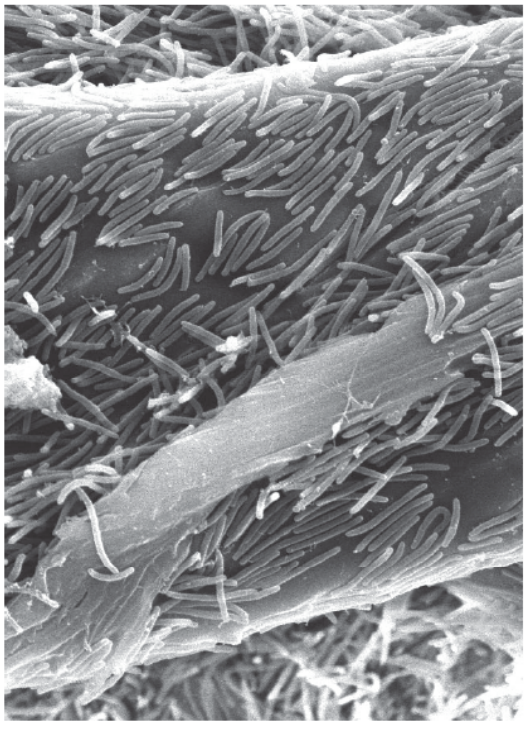
1.3 Biology of *C. hutchinsonii*

C. hutchinsonii was first isolated from soil and a description was published by Hutchinson and Clayton under the name '*Spirochaeta cytophaga*' in 1919

[Hutchinson & Clayton 1919]. In the early days of cellulose decomposition research, this bacterium was one of the major species studied by scientists. The complete genome sequence of *C. hutchinsonii* has now been published [Xie et al. 2007]. A total of 3,790 protein-encoding genes are predicted, 1,986 of which have been assigned a tentative function, and the GC content is 38.85 % [Xie et al. 2007]. *C. hutchinsonii* grows as flexible rods, with slightly tapering ends. Cells are 0.3 – 0.5 × 2 – 10 μm, usually 0.4 × 2 – 5 μm. The genus *Cytophaga* and similar bacteria are well known for their rapid gliding motility [McBride 2004]. On agar surfaces, the cells sometimes show very fast gliding, suddenly jumping forward with intense wriggling, although only a few cell lengths at a time [Jahn & Bovee 1969]. Some proteins related to gliding motility have been studied in *Cytophaga johnsonii* (Agarwal et al. 1997, McBride & Baker 1996, McBride 2001, McBride 2004). The cells contain a bright yellow pigment, which becomes deep brown-red after being covered by 20 % KOH [Larkin 1989]. The yellowish substance shows highest adsorption at 450 nm which suggests it is a carotenoid [Verma & Martin 1967].

C. hutchinsonii is capable of completely degrading crystalline cellulose (Figure 1-8) [Larkin 1989, Rubin 2008, Xie et al. 2007]; however, its complete enzymatic system for cellulose degradation is still unknown, though the genome of *C. hutchinsonii* has been sequenced. Louime et al. (2006) reported that almost 70 % of CMCase activities were in the periplasm of *C. hutchinsonii*. One of the endoglucanases (CHU_1336) from *C. hutchinsonii* has been cloned and over-expressed in *E. coli*, and was repeated to be highly active against carboxymethyl cellulose (CMC) [Louime et al. 2007]. The glycoside hydrolase (GH) complement in *C. hutchinsonii* includes 48 enzymes belonging to 18 families. Candidate cellulases

belong to GH families 1, 3, 5 and 9. Interestingly, no exoglucanase has been found in the whole genome [Wei et al. 2009, Wilson 2008, Xie et al. 2007]. The subcellular locations of these cellulases are not known. Automatic prediction of locations for these cellulases shows various results suggesting that these cellulases may be located in the membrane system. Xylanase and mannanase are also found, though *C. hutchinsonii* is reported not to use xylan and mannose as sole carbon sources [Larkin 1989, Xie et al. 2007]. Insertion of plasmids directly to *C. hutchinsonii* has been reported by Xu et al. (2011). McBride and Baker (1996) have reported transformation of *C hutchinsonii* by conjugation. Direct adhesion is needed for *C. hutchinsonii* to hydrolyze cellulose, but no obvious encoded protein has been found relating to adhesion [Xie et al. 2007]. This remains a question to be understood.



(Source: Xie et al. 2007)

Figure 1-8. Scanning electron micrograph of *C. hutchinsonii* cells digesting cellulose filter paper.

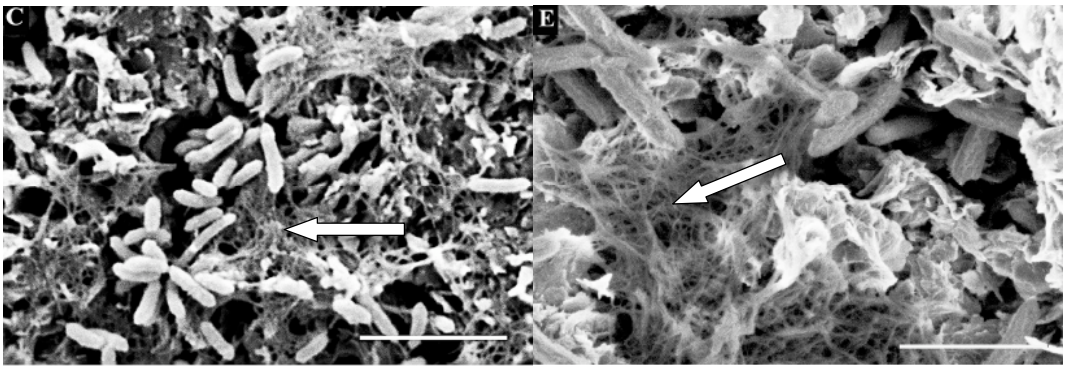
At least 16 enrichment or isolation media recipes have been described for *C. hutchinsonii* or related species [Atlas 1997]. Mineral salt related media are suitable for isolation of *C. hutchinsonii*. Colonies on filter paper are large, spreading and bright yellow, the center soon becoming translucent and somewhat slimy. Colonies on glucose mineral agar are bright yellow, more or less compact to slightly spreading, with an entire or wavy edge and moderately raised to flat [Larkin 1989]. In liquid media with cellulose or glucose, slime may be produced, making the culture viscous and harvesting difficult. Diethyleneglycol is reported to be a good solvent for slime isolation [Verma & Martin 1967]. *C. hutchinsonii* can use peptone, yeast extract, several amino acids, nitrate and ammonium as nitrogen source. Cellulose [McBride & Baker 1996, Walker & Warren 1938], CMC [Louime et al. 2006], cellodextrin [Zhu et al. 2010], cellobiose, or glucose can be used as sole carbon source [Larkin 1989].

C. hutchinsonii produces and releases slimy substances when cultured in liquid, making the medium viscous. Slime production in *C. johnsonii* was reported to increase with culture maturity [Follett & Webley 1965]. The slime is reported to consist of polysaccharides composed of glucose, mannose, arabinose, xylose and glucuronic acid residues [Martin et al. 1968].

1.4 Microbial extracellular polysaccharide

Microbial extracellular polysaccharides (EPS) (Figure 1-9) are synthesized by microorganisms and secreted to the cell surface or the ambient environment [Paul 2008]. The abbreviation ‘EPS’ is variously used to mean ‘extracellular polysaccharides’, ‘exopolysaccharides’, ‘exopolymers’ or ‘extracellular polymeric

substances'. However, 'extracellular polymeric substances' may cover a much wider range of materials, including proteins and nucleic acids [Wingender et al. 1999]. The most common synthetic pathway of EPS occurs at the cell membrane or in the periplasm [Harrah et al. 2006]. The release of EPS to the extracellular environment occurs through the processes of active secretion, shedding of cell surface material, cell lysis and adsorption from the environment [Wingender et al. 1999]. Sutherland (1985) suggests that wall polymers may either be secreted or be lost from the Gram-negative cell surface, and that the distinction between extracellular and wall polysaccharides can only be approximate.



(Source: Rollefson et al. 2011)

Figure 1-9. SEM images of biofilms from *Geobacter sulfurreducens*. A SEM image (left) and a field emission SEM image (right) of *G. sulfurreducens*. Scale bar is 5 μm in the left image and 2 μm in the right image. Both arrows indicate the EPS produced by *G. sulfurreducens*.

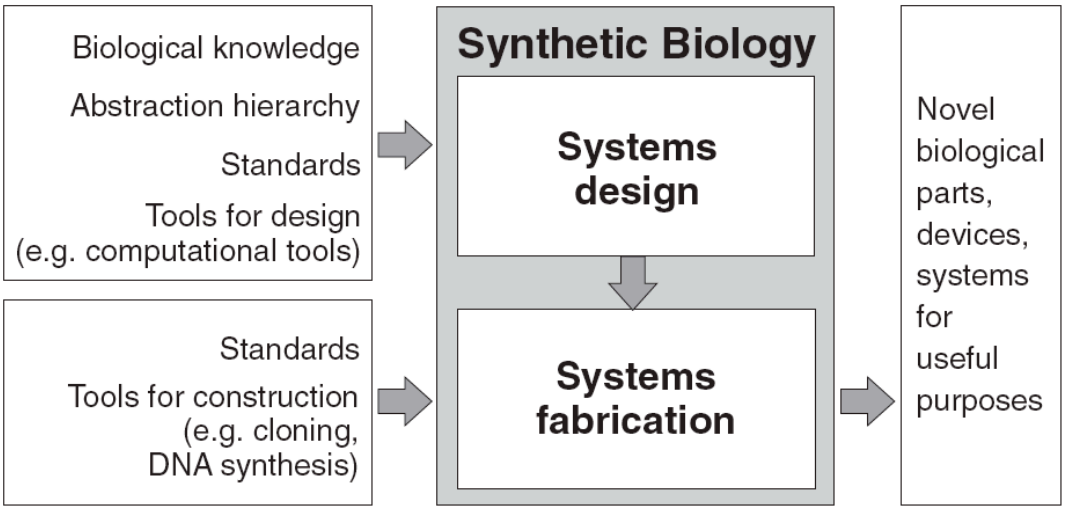
The term 'glycocalyx' was introduced by Costerton to refer to all EPS-containing structures, such as capsules, sheaths and slimes [Wingender et al. 1999], but Sutherland (1990) states that 'glycocalyx' is unsatisfactory to differentiate chemical entities found on the surface of cells. Two different types of EPS were defined by Nielsen and Jahn (1999): the bound and the soluble EPS. Sheaths, capsular polymers, condensed gel, loosely bound polymers and attached organic

material are included in the bound EPS, and soluble macromolecules, colloids and slimes are included in the soluble EPS. Most microbial EPS (apart from capsules) are highly soluble in water or dilute salt solutions [Sutherland 1999], and different species of bacteria may produce similar ranges of EPS [Sutherland 1997]. The EPS in many Gram-negative bacteria are relatively simple, and may be either homopolysaccharides, which are usually polymers composed of D-glucose, or heteropolysaccharides, which consist of 2 to 4 types of monosaccharide residues, sometimes acylated with regular repeating units ranging in size from disaccharides to octasaccharides [Sutherland 2001]. D-Glucose, D-galactose, D-mannose (in the pyranose forms), L-fucose and L-rhamnose are frequently found in microbial EPS [Sutherland 1990]. Hernandez-Mena and Friend (1993) identified glucose, mannose, galactose, rhamnose, fucose and glucuronic acid from a mixed microbial slime. A similar EPS composition was also identified by Jiao et al. (2010) from acidophilic microbial biofilms. Eukaryotic polysaccharides may contain pentoses such as D-ribose or D-xylose, but these are of less common occurrence in extracellular polymers derived from prokaryotes, [Sutherland 1990] except Cyanobacteria, in which D-xylose may occur [Plude et al. 1991, Sutherland 1990]. The ketals or uronic acids in EPS of Gram-negative bacteria are found in linear polyanionic macromolecules [Sutherland 2001]. D-Glucuronic acid is the most common uronic acid found in microbial EPS, where D-galacturonic acid is less common and D-mannuronic acid is found in very few species [Sutherland 1990].

Cell adhesion to surfaces is one important function of EPS. This can aid the colonization of an inert or tissue surface or accumulation of bacterial cells on nutrient-rich surface in oligotrophic environments [Wingender et al. 1999]. The EPS

produced by *Flexibacter* BH3 exhibits viscous properties and increases adhesion ability [Humphrey et al. 1979]. It has also been suggested that EPS is positively correlated to surface adhesion from an active sludge analysis [Cammarota & Sant'Anna Jr. 1998]. Tsuneda et al. (2003) analyzed 27 bacterial strains from a wastewater treatment reactor and found that EPS-producing strains promoted cell adhesion. In addition to cell adhesion, EPS is also functional in cell-cell interaction, environmental stress response and formation of a physical protective barrier [Weiner et al. 1995]. EPS also plays a major role in heavy metal adsorption and bioremediation (bioleaching). Results of EPS analysis from *Hymenobacter aerophilus* showed that EPS was correlated to an increase in the number and type of functional groups on the cell surface and was positively correlated to metal adsorption [Baker et al. 2010]. Bhaskar and Bhosle (2006) evaluated the relationship between the physical environment and the heavy metal-binding ability of EPS from a marine bacterium of the genus *Marinobacter*. Their results showed that EPS absorb more Cu^{2+} and Pb^{2+} at neutral pH than at acidic pH.

Microbial EPS has a wide range of applications in the food industry, as well as non-food uses, in pharmaceuticals and cosmetics, and in the oil industry [Sutherland 1996]. EPS synthesized by lactic acid bacteria plays a major role in the manufacture of fermented dairy products, such as yoghurt, drinking yoghurt, cheese, fermented cream, and milk-based desserts [Duboc & Mollet 2001]. Microbial EPS, such as xanthan, gellan, pullulan, ß-D-glucans and hyaluronic acid are commercially important [Sutherland 1998].



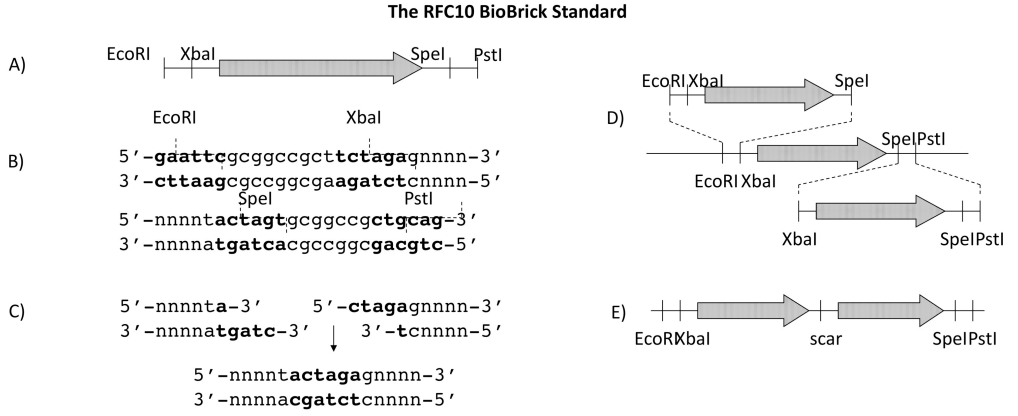
(Source: Heinemann and Panke, 2009)

Figure 1-10. Scheme of Synthetic biology. Synthetic biology encompasses systems design and fabrication. Each part has its specific prerequisites and inputs. Ultimately, synthetic biology will deliver novel biological entities with improved functionality.

1.5 Synthetic biology

Synthetic biology merges biological, engineering and chemical knowledge and has already led to a wide range of achievements in science and technology development [Haseloff & Ajioka 2009] (Figure 1-10). Scientists use biotechnological skills and engineering concepts to make controllable biological parts or devices turning unnatural or artificial parts into new functional systems [Andrianantoandro et al. 2006, Mukherji & Oudenaarden 2009]. Products and processes, such as chemicals, drugs, biosensors, biofuels, and bioremediation represent some new achievements in this field [P.O.S.T. 2008]. During the last thirty years, many experiments and much laboratory work have solidified protocols and theories in modern biochemistry and molecular biology. Novel technology-dependent sciences such as bioinformatics and system biology are also enabling scientists to establish the foundations of synthetic biology, supporting interdisciplinary research, such as BioBrick biological part development, which was first proposed by Dr. Thomas

Knight in 2003. In the essential concept of the BioBrick biological part, a library of DNA inserts are prepared in a specially designed vector flanked by four crucial restriction enzymes: EcoRI and XbaI at the upstream end, and SpeI and PstI at the downstream end, as shown in Figure 1-11. The BioBrick assembly standard enables the distributed production of a collection of compatible biological parts. Since engineers carry out the exact same operation every time that they want to combine two BioBrick parts, the assembly process is amenable to optimization and automation, in contrast to more traditional ad hoc molecular cloning approaches [Shetty et al. 2008].



(Source: French 2009)

Figure 1-11. Basic design of BioBricks. (a) Each BioBrick is a length of DNA bearing a genetic component such as an open reading frame, ribosome-binding site, promoter, transcription termination sequence or any combination of these. Each BioBrick possesses EcoRI and XbaI restriction sites at the 5' end, and SpeI and PstI sites at the 3' end. (b) Standard prefix and suffix sequences for BioBricks. The six-base pair recognition sites for each restriction endonuclease are shown in bold and dashed lines indicate the staggered cuts made by each enzyme. (c) Ligation of an SpeI-cut end to an XbaI-cut end generates a six-base pair 'scar', which is not recognized by either XbaI or SpeI. (d) By appropriate choice of restriction enzymes, any BioBrick can be inserted either upstream or downstream of any other BioBrick. (e) In either case, the product, bearing both components, is also a BioBrick, bearing the same four restriction sites as the original component BioBricks. It can thus be added either upstream or downstream of any other BioBrick. In this way large and complex constructs can be built up quickly and easily from a library of standard parts. [French 2009]

1.6 Aims of this project

Cellulose is important for the biofuel industry. Attempts to transfer cellulose-degrading systems to heterologous hosts have had very limited success so far [French 2009]. Novel biotechnology research based on the concepts of Synthetic Biology, such as BioBrick assembly, is potentially very useful to combine DNA sequences and gererate large numbers of combinations easily, therefore could be applied to test simultaneous expression of multiple proteins. This can be applied to study the role of synergy in enzymatic cellulose degradation. Therefore, one main aim of this project was to establish a basic methodology for testing large numbers of cellulases from genome sequences through the mechanism of BioBrick parts assembly, in a form which may then allow them to be tested in combinations for synergistic effects.

Although *C. hutchinsonii* was discovered almost ninety years ago and is well known for its cellulose degradation ability, very few studies have focused on this bacterium. Also, since the genome sequence was published in 2007, only CHU_1336 [Louime et al. 2007] had been characterized. The entire cellulase system in this bacterium is still unknown. In order to analyze the cellulases in *C. hutchinsonii*, one basic strategy to identify those possible cellulases is by bioinformatics. Candidate genes can then be cloned as BioBricks, which will later allow them to be easily combined. Once the expression has been achieved, the analysis of each cellulase candidate must consider multiple possible activities in case they are incorrectly annotated or may have multiple activities. Clones showing activity can then be tested in combinations, facilitated by flexible BioBrick assembly. Once this strategy has been validated in *C. hutchinsonii*, and any problems have been identified and fixed,

larger scale application to many other genome sequences and metagenomic sequence sets could be applied and analyzed, moving the assays to high throughput formats.

Hence, another aim of this project was to verify the putative cellulases and other polysaccharide hydrolysis-related enzymes suggested by bioinformatic analysis. With the bioinformatics analysis, this project also attempted to cast light on the cellulose degradation mechanisms of *C. hutchinsonii*.

Following the results of bioinformatics analysis, more cellulase candidates would be characterized for cellulose hydrolyzing ability. The cellulase candidates would be expressed in different host strains. Assays were established to test the expressed cellulases and to identify their cellulose degrading abilities. The cellulase characterization of *C. hutchinsonii* also aimed to determine whether there was any exoglucanase, which is reported to be absent in its genome sequence.

Since degradation of cellulose by C. hutchinsonii produces slime (EPS), the composition of this slime is potentially important since EPS could be useful for industrial utilization. The only research on the EPS of C. hutchinsonii was in 1968 by Verma and Martin, and this only tested the EPS produced from glucose-grown culture. Therefore this project attempted to analyze EPS from different cultures grown on different carbon sources.

Chapter 2 Materials and Methods

Chemicals, reagents, buffers and bacterial culture media applied in this research are listed in Appendix I. Bacterial strains used in this research are listed in Table 2-1.

Table 2-1. List of bacteria strains

Strain	Genotype	Description
C. hutchinsonii		Strain ATCC 33406 (Obtained from DSMZ, strain No. 1761)
Escherichia coli JM109	*endA1*, *glnV44*, *thi-1*, *relA1*, *gyrA96*, *recA1*, *mcrB*$^+$, Δ(*lac-proAB*), *e14-*, [*F' traD36 proAB*$^+$ *lacI*^{q} *lacZ*Δ*M15*] *hsdR17*(r_K^- m_K^+)	Obtained from Promega. (Catalog No. P9751)
E. coli Rosetta	F^-, *ompT*, *hsdS*$_B$(R_B^- m_B^-), *gal*, *dcm* λ(DE3 [*lacI lacUV5-T7* gene 1 *ind1 sam7 nin5*]), *pLysSRARE* (*Cm*^{R})	Obtained from Novagen (Catalog No. 70956-3). The Novagen competent cell Rosetta 2 (DE3)pLysS host strain is BL21 derivative designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*.
E. coli JW2120-1	F^-, Δ(*araD-araB*)*567*, Δ*lacZ4787*(::*rrnB-3*), λ^-, Δ*bglX758*::*kan*, *rph-1*, Δ(*rhaD-rhaB*)*568*, *hsdR514*	The *bglX* mutant strain. Obtained from the Keio Knockout Collection of single-gene knockouts.
E. coli BL21(DE3)pLysS	F^-, *ompT*, *gal*, *dcm*, *lon*, *hsdS*$_B$(r_B- m_B-), λ(DE3), *pLysS*(*Cm*^{R})	Obtained from Promega. (Catalog No. L1191)
C. freundii NCIMB11490		ATCC 8090, Obtained from NCIMB, Aberdeen, U.K..

2.1 Cultures, TLC, Colorimetric assay and HPLC analysis

2.1.1 Culture of *C. hutchinsonii*

The mineral medium used for *C. hutchinsonii* culture was modified from Dubos Salts Medium [Atlas 2010]. Cells were cultured on DSM3T agar plates (Appendix I) supplemented with carbon sources: 0.2 g glucose, 0.2 g cellobiose or 0.2 g filter paper strip. The plates were incubated at 30 ºC and checked daily. For liquid cultures, cells were cultured in 100 ml of DSM3T medium (Appendix I) incubated at 28 ºC, shaking at 100 rpm for 14 days supplemented with carbon source: 0.5 g filter paper, 0.5 g glucose or 1.0 g cellobiose. Filter paper was autoclaved, glucose and cellobiose were filtered (0.22 μm) before use. Another culture was performed with 0.2 g avicel in 100 ml DSM3T medium and incubated at 28 ºC, shaking at 100 rpm for 8 days. Each carbon source was tested in triplicate. One millilitre of each culture was sampled daily and its optical density (OD_{600}) was measured (except day 1 and 13). The one millilitre sample was then preserved in the freezer (-20 ºC) for further use.

Samples shown in Figure 3-9 were from a 100 ml DSM3T culture with 0.2 g filter paper. This culture was incubated at 100 rpm, 30 ºC for 14 days. The filtrate (0.22 μm filter) was kept for TLC analysis or further acid digestion as described in Section 2.1.2.2

2.1.2 TLC analysis

2.1.2.1 Sample preparation for TLC

The frozen samples (1 ml) described in Section 2.1.1 were thawed for analysis by thin layer chomatography (TLC). Before loading to the TLC plate,

samples were centrifuged at 11000 *g* for 10 min. The supernatant was collected and then directly used for TLC sample loading as described later.

2.1.2.2 Sample digestion with hydrochloric acid

The acid digestion was based on a previous report [Sturgeon 2000]. Hydrochloric acid (1 M final concentration) and 1.1 ml sample were mixed and incubated in a 90 ºC water bath for 3 h. The mixture was cooled down and neutralized (about pH 7.0) with sodium hydroxide before being used for TLC.

2.1.2.3 Procedures for TLC

The TLC was performed according to instruction by Dr. Amjad Iqbal and the protocols were from Fry (1988). TLC was performed on 0.2 mm thick, pre-coated silica gel-60 plastic sheets without fluorescent indicator (obtained from Merck). Mobile phase was prepared by mixing 1-butanol : acetic acid : water in the ratio 2 : 1 : 1 by volume. Spot detection reagent was prepared by mixing 0.5 g thymol, 5 ml sulfuric acid and 95 ml ethanol [Jork et al. 1990, 1994]. Arabinose, cellobiose, fructose, fucose, galactose, glucose, lactose, maltose, mannose, raffinose, rhamnose, sucrose and xylose were prepared at 1 mg/ml as standards. Appropriate volumes of sample and standards were transferred to the TLC plate at 0.5 µl per dot. The TLC plate was then dried before being carefully put into a glass chamber with 20 ml mobile phase. After 5 h development, the plate was removed from the glass chamber and dried in the fume hood by ambient air. It was then immersed into thymol reagent

for 3 sec and dried again with ambient air in the fume hood. Finally, the plate was heated at 110 ºC for 10 min and examined visually.

2.1.3 Acid digestion of filter paper

The acid digestion with sulfuric acid basically followed the protocols of Fry (1988). Pure filter paper was digested by acid as a reference for colorimetric assays described in Chapter 3. An autoclaved and dried piece of filter paper (50 mg) was shredded into small fragments. These fragments were transferred into a 150 mm height test tube with addition of 1 ml 72 % sulfuric acid. The test tube was kept cool in an ice bath and gently swirled by hand with occasional shaking until all paper fragments were entirely dissolved. After 1 h, 2 ml water was added and mixed with a quick swirl leading to a final concentration of about 33 % sulfuric acid in total. The test tube was then immersed in a boiling water bath for 40 min. The test tube was then cooled down quickly and another 7 ml of water was added to dilute its acidity. The resulting liquid was then directly applied for hexose colorimetric assay.

2.1.4 Colorimetric assays for metabolic product determination

Colorimetric assays are widely used for quantification of monosaccharides, deoxy-sugars and uronic acids. The general concept is based on forming furfural or related chemicals in strong acids, sometimes with heating [Dische 1962]. In this research, hexose, pentose and uronic acid content were tested by thee different assays as described below.

2.1.4.1 Anthrone-sulfuric acid method for total hexose quantification

The procedure for the anthrone method was modified from published procedures [Loewus 1952, Ludwig & Goldberg 1956, Morris 1948]. The advantages of this revised procedure are to reduce the usage volume of sulfuric acid and to use no ethyl acetate. Each test tube was prepared with 0.5 ml of diluted sample or glucose standard (0, 2, 10, 20, 50, 100 µg/mL) and then gently loaded with 1 ml of 2 % (w/v) anthrone-sulfuric acid reagent. After vigorous swirling, the test tube was placed into an ice bath immediately to cool down. A range of colors were produced, from bright yellow at low concentration of sugar to dark green at high concentration. The absorbance was measured at 620 nm against de-ionized water. Data for the glucose standard regression equation was determined in triplicate at each concentration.

2.1.4.2 Pentose assay protocols and standard curve construction

The pentose assay used in this research was the method of Bial [Dische 1962, Fry 1988]. This assay is widely used for determinations of D-ribose nucleotides and D-ribose nucleic acids [Dische 1962]. The test was performed by conversion of pentose into furfural in the presence of hot acid, which then reacts with orcinol to give green colors [Nigam and Ayyagari 2007].

The sample volume for Bial's pentose assay was 0.5 ml and this was prepared in the test tube. Orcinol was prepared as 6 % (w/v) solution in ethanol and 67 µl was added into the test tube. Another 1.0 ml of ferric chloride [0.1 % (w/v) in concentrated hydrochloric acid] was also added into the test tube and mixed well

with swirling. The mixture was heated in a 100 ºC water bath for 20 min and then cooled down to room temperature. The resulting green color was measured at 665 nm. Xylose was used for the pentose standard and was prepared at five concentrations: 2, 4, 8, 16 and 20 µg/ml. Each concentration was tested in triplicate and the regression equation was calculated from these triplicate data.

2.1.4.3 Uronic assay protocols and standard curve construction

Uronic acid assay was directly applied from Fry (1988) and Chaplin (1986). Sample (0.2 ml) was prepared in a test tube and 1 ml sodium tetraborate decahydrate [0.5 % (w/v) in concentrated sulfuric acid] was added with gentle mixing. The test tube was then heated in a 100 ºC water bath for 5 min and then cooled down to room temperature. Absorbance was measured at 520 nm. Next, 20 µl of 0.15 % (w/v) 3-hydroxydiphenyl, which was prepared in 1 M sodium hydroxide, was added into the test tube with thorough mixing. After 5 min at room temperature (25 ºC), the absorbance was measured again at 520 nm. The result was calculated from the increase of the absorbance. Glucuronic acid was used as the standard uronic acid and was prepared at five concentrations: 5, 25, 50, 75 and 100 µg/ml. Each concentration was tested in triplicate for the regression equation.

2.1.5 Determination of metabolic products by HPLC

Analysis of the metabolic products was attempted by high-performance liquid chromatography (HPLC). Two HPLC systems were applied. In the French laboratory, the supernatant from a 100 ml DSM3T filter paper (0.2 g) culture was first passed through a 0.22 µm filter. The filtrate was then hydrolyzed by

hydrochloric acid as described in Section 2.1.2.2. The HPLC column in Dr. French's laboratory was LC-NH_2 5 μm (25 cm × 4.6 mm, Supelco) and the samples were analyzed using a refractive index detector (Gilson, Model 131). The mobile phase was water (25 %)/acetonitrile (75 %) with de-gassing before use. The sample was run at room temperature with flow rate 0.5 ml/min.

The acid-digested sample was also analyzed in Prof. Fry's laboratory by Mr. Tim Gregson. The detction was using the Dionex Ion Chromatography System with autosampling (Model AS-3500). Samples were run through a Dionex Carbopac PA-1 column (4 x 250 mm) with integrated pulsed amperometric detection (Dionex ED40 Detector). The mobile phase was operated in a gradient program as described in Table 2-2. Samples were monitored and analyzed using the Chromeleon 6.8 Chromatography Data System software package from Dionex.

Table 2-2. Program of mobile phase in Dionex HPLC

Time (min)	Chemical	Elution Method
0 ~ 3	20 mM NaOH	isocratic
3 ~ 44.5	100 % Water	isocratic
44.5 ~ 75	100 % Water – 800 mM NaOH	linear gradient
75 ~ 81	800 mM NaOH	isocratic
81 ~ 82	800 mM NaOH – 20 mM NaOH	linear gradient
82 ~ 90	20 mM NaOH	isocratic

2.2 Bradford method for protein concentration determination

All protein quantification results were determined using the Quick Start Bradford protein assay (obtained from Bio-Rad). A standard curve was established

by detecting 2, 4, 6, 8 and 10 µl of 2 mg/ml BSA with 1 ml Bradford reagent. The final linear regression equation was calculated from triplicate measurements. Unknown samples were prepared in 20 µl and mixed with 1 ml Bradford reagent. All assays were mixed well in cuvettes by inversion and absorbance was measured at 595 nm within 1 h.

2.3 Plasmid construction for expression

2.3.1 DNA cloning

Polymerase chain reaction (PCR) was applied for gene, DNA amplification and mutagenesis. KOD Hot Start DNA polymerase (Novagen) was chosen for PCR. All components and conditions for PCR are shown in Tables 2-3 and 2-4. The details of all primers for PCR are listed in Appendix II.

Table 2-3. Components for KOD polymerase PCR

Component	Volume	Final Concentration
10X Buffer for KOD Hot Start DNA Polymerase	5 µl	1X
25 mM $MgSO_4$[a]	3 µl	1.5 mM
dNTPs (2 mM each)	5 µl	0.2 mM (each)
PCR Grade Water	X µl	
Sense (5') Primer (10 µM)	1.5 µl	0.3 µM
Anti-Sense (3') Primer (10 µM)	1.5 µl	0.3 µM
Template DNA[b]	Y µl	
KOD Hot Start DNA Polymerase (1 U/µl)	1 µl	0.02U/µl
Total reaction volume	**50 µl**	

(Novagen)

Table 2-4. Conditions for KOD polymerase PCR

Step	Target size			
	< 500 bp	500–1000 bp	1000–3000 bp	> 3000 bp
1. Polymerase activation	95°C for 2 min	95°C for 2 min	95°C for 2 min	95°C for 2 min
2. Denature	95°C for 20 s	95°C for 20 s	95°C for 20 s	95°C for 20 s
3. Annealing	Lowest Primer Tm°C for 10 s			
4. Extension	70°C for 10 s/kb	70°C for 15 s/kb	70°C for 20 s/kb	70°C for 25 s/kb
Repeat steps 2–4	20–40 cycles. For more information see "Cycle number" below			

(Novagen)

All cellulase genes used in Chapter 5 and Chapter 6 were directly cloned from *C. hutchinsonii* cells. Cells were taken from an agar plate and suspended in de-ionized water as the template for PCR. PCR products were checked by agarose gel electrophoresis, staining with GelGreen stain (Biotium). The successful PCR products were purified by adsorption to glass beads (Appendix III) and then checked again by agarose gel electrophoresis.

The purified DNA was ligated with SOB linear vector. SOB linear vector was modified from pSB1A2 (Registry of Standard Biological Parts) which includes four BioBrick restriction endonuclease recognition sites: PstI, EcoRI, XbaI, SpeI. The detailed ligation protocol is listed in Appendix III. This blunt-end ligation was then used as the template for fusion PCR. Fusion PCR aimed to amplify the correct DNA sequence from a blunt-end ligation; therefore, one primer from the target cellulase and one primer from the SOB vector were used in this PCR. Fusion PCR products were checked by agarose gel electrophoresis and purified using glass beads prior to the blunt-end self-ligation. Some fusion PCR products, especially the products with too many un-wanted bands shown on the gel electrophoresis, were excised and purified from agarose gels using SYBR-Safe stain (Invitrogen) under blue light before the blunt-end self-ligation.

The blunt-end self-ligation product was transformed into *E. coli* JM109 competent cells (Appendix III). The pure cellulase DNA sequence cloned in the plasmid was harvested by plasmid DNA miniprep (Appendix III). The plasmids were checked for correct construction by two methods: restriction endonuclease digestion with gel electrophoresis, and DNA sequencing.

2.3.2 Mutagenesis

RFC10 BioBricks may not possess EcoRI, XbaI, SpeI and PstI sites (Figure 2-1) based on the RFC10 BioBrick standard [Knight 2003]. Thus, all these four restriction endonuclease cutting sites in target DNA sequence were mutated by PCR. Primers for mutagenesis were designed and are listed in Appendix II. The mutated PCR product was purified and self-ligated. If one DNA sequence had several sites to be mutated, the self-ligation product was used as the template for another mutagenic PCR. This 'PCR→purification→self-ligation' procedure was repeated until all BioBrick forbidden sites were mutated. Then the final self-ligation product was transformed into *E. coli* JM109 and the pure plasmids were prepared by miniprep. The mutagenesis was confirmed by DNA sequencing.

2.3.3 Ribosome binding site insertion

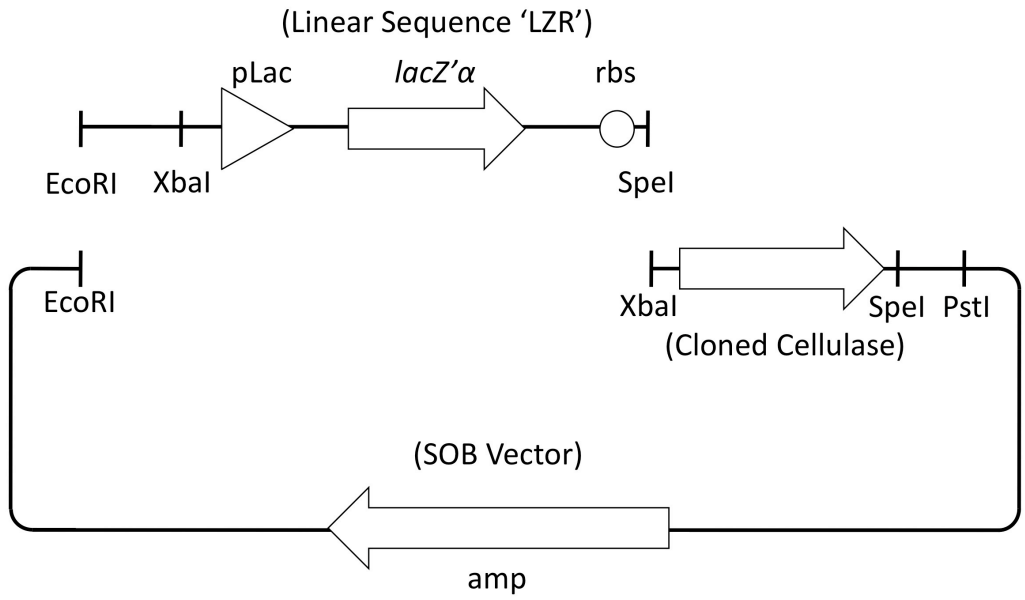
A BioBrick-designed linear DNA sequence – 'LZR' – containing a *lac* promoter, *lacZ'α* and ribosome binding site (rbs) was prepared by PCR. Following the BioBrick parts assembly rule, the LZR insertion was designed as shown in Figure 2-1(a) where LZR was inserted upstream of each target gene. The LZR was digested with EcoRI and SpeI, and the plasmid was digested with EcoRI and XbaI. These two

fractions were ligated and formed a new BioBrick construct [Figure 2-1(b)]. The compatible XbaI and SpeI ends were ligated, leading to a "scar" sequence nor recognized by either enzyme. This ligation product was transformed into *E. coli* JM109 and cells were grown on blue/white screening plates. The blue transformants were picked for plasmid DNA miniprep. The correct plasmids were confirmed by DNA sequencing.

2.3.4 Expression vector construction in pT7-7

Four predicted GH family 3 ß-glucosidases were cloned and inserted in pT7-7 vector. Primers (Appendix II) were designed according to the restriction endonuclease cutting sites from pT7-7. The predicted N-terminal signal peptide was eliminated when designing these primers. PCR was first performed with KOD polymerase and checked by agarose gel electrophoresis. After purification, the linear PCR product and pT7-7 vector were digested by appropriate restriction endonucleases as shown in Table 2-5. The digestion was stopped by inactivating the enzyme at the appropriate temperature (65 ºC or 80 ºC) for 20 min. This product was then ligated overnight at 16 ºC. Ligation was stopped by heating at 65 ºC for 20 min prior to being transformed into *E. coli* JM109. The plasmids were then purified by plasmid DNA miniprep and confirmed by agarose gel electrophoresis and DNA sequencing.

(a)



(b)

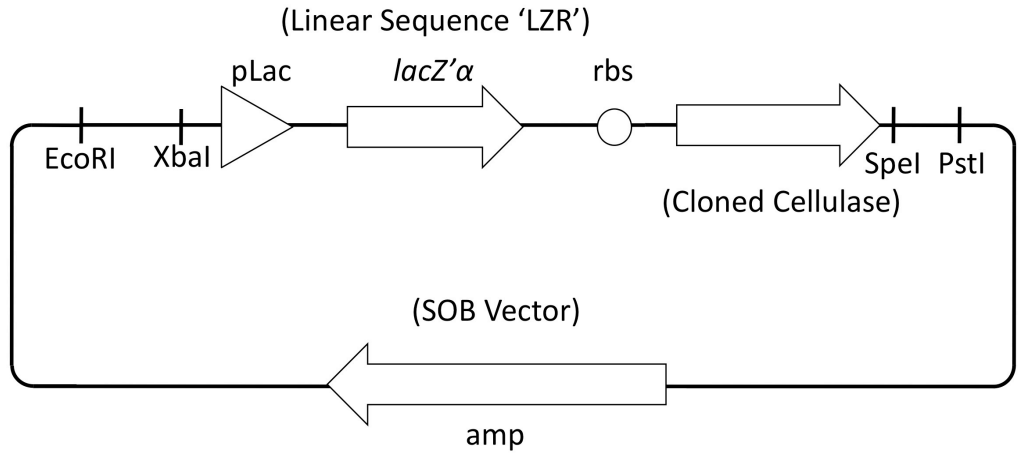


Figure 2-1. Diagram of rbs insertion. (a) BioBrick constructed vector is digested with EcoRI and XbaI. The BioBrick prefix is present upstream of linear 'LZR' and SpeI is located downstream of LZR. LZR is digested with EcoRI and SpeI. (b) When LZR is inserted; the BioBrick prefix is intact and still includes EcoRI and XbaI sites. The compatible part of SpeI and XbaI are ligated between the rbs and 5' start of the cloned cellulase DNA sequence, and the ligated product is no longer recognized by either XbaI or SpeI.

Table 2-5. Restriction endonuclease sites used for insertion into pT7-7

Cellulase Name	Restriction Endonuclease Sites
Chu_2268	NdeI, PstI
Chu_2273	EcoRI, PstI
Chu_3577	NdeI, EcoRI
Chu_3784	EcoRI, PstI

Table 2-6. Plates used for cellulase transformation and expression

Bacterium strain	Agar Plate
E. coli JM109	LB/amp/1% glucose
E. coli JW2120-1	LB/amp/1% glucose
E. coli Rosetta	LB/amp/cam/1% glucose
E. coli BL21(DE3)pLysS	LB/amp/cam/1% glucose
C. freundii	LB/carb/1% glucose

[Note: amp: ampicillin; cam: chloramphenicol; carb: carbenicillin.]

2.4 Enzyme Expression

The target cellulases were expressed in *E. coli* JM109, *E. coli* JW2120-1, *E. coli* Rosetta and *C. freundii*. Cellulase plasmids constructed with pT7-7 were transformed and expressed in *E. coli* BL21(DE3)pLysS. The plates applied for each strain are listed in Table 2-6. Additional glucose in plates was intended to prevent minor expression due to leaking of the *lac* promoter.

2.4.1 Expression in LB medium

Randomly-selected colonies on transformation plates were subcultured on a new plate as indicated in Table 2-6. PCR was applied to confirm that plasmids were harbored in subcultured cells. The positive subcultures were then used for expression. The initial culture was prepared in 5 ml of LB with appropriate antibiotics incubating at 37 ºC overnight. A larger culture of 50 ml LB with appropriate antibiotics and 1 ml of the initial culture was prepared next day. This flask was incubated at 37 ºC, 180 rpm and IPTG was added (final concentration 0.38 mM) after OD_{600} reached 0.2 (about 2 h). For 1 l expression in *E. coli* BL21(DE3)pLysS, 50 ml LB with ampicillin (80 µg/ml) and chloramphenicol (40 µg/ml) was incubated overnight at 37 ºC. This culture was then divided evenly and transferred into two 2-litre flasks with 500 ml LB in each flask. The flasks were incubated at 37 ºC for 4 h and then induced with IPTG. The IPTG induction of *E. coli* JM109, *E. coli* JW2120-1, *E. coli* Rosetta and *C. freundii* used 0.38 mM IPTG for 6 h. For both 50 ml and 1 l expressions in *E. coli* BL21(DE3)pLysS, the induction was with 0.07 mM IPTG for 5 h.

2.4.2 Expression in M9 medium

M9 medium was also used for expression in *E. coli* JM109 when expression was toxic. The M9 recipe is given in Appendix I. The expression culture was prepared in 50 ml M9 with appropriate antibiotics which was supplemented daily. Glycerol (1 %) was added as the carbon source. The culture was grown at 20 ºC for 2

days and the expression was induced with 0.15 mM IPTG for 24 h when OD_{600} was about 0.2.

2.4.3 Cell lysis

After cellulases were expressed, cells were harvested by centrifugation at 2000 *g* for 15 min. The pellet was re-suspended in 1 ml PBS (pH 7.0±0.2), and then transferred to a 1.5 ml micro-centrifuge tube in an ice bath. Cell lysis was prepared by using an ultrasonic disintegrator (Soniprep 150 Sanyo/MSE) at 12000 Hz with 10 s breaking and 40 s rest in an ice slurry for 10 cycles. The mixture was centrifuged at 11000 *g* for 2 min. The crude cell lysate (supernatant) was transferred to a new 1.5 ml micro-centrifuge tube and the cell debris (pellet) was re-suspended in 1 ml PBS (pH 7.0±0.2). All crude lysate and cell debris were preserved at -20 ºC for further use.

2.4.4 Expression toxicity tests

The cellulase expression toxicity in *E. coli* JM109 was tested. For each cellulase, a starter culture (5 ml LB) with appropriate antibiotics was incubated at 37 ºC overnight. The next day, two flasks (50 ml LB each) with appropriate antibiotics and 1 ml starter culture were incubated at 37 ºC. After 2 h incubation, 0.38 mM IPTG was added to one flask. Both the cultures were incubated for another 4 h. During the total 6 h incubation, optical density (OD_{600}) was recorded every 1 h for both cultures. For each cellulase, the culture was tested in triplicate.

For the cellulase expression toxicity test in *E. coli* Rosetta, *E. coli* JW2120-1 and *C. freundii*, all preparations were the same as in *E. coli* JM109 except that OD_{600}

was measured only at the final hour. Also, the induction time duration was 6 h. Each test was also performed in triplicate. The protocols of expression toxicity tests in *E. coli* BL21(DE3)pLysS were almost the same as in *E. coli* JM109 except that the IPTG was added in different concentrations (0, 0.02, 0.04, 0.23 and 0.38 mM) for 4 h.

2.4.5 Cellobiohydrolase and ß-glucosidase activity assay

The crude lysate (20 µl) was mixed with 960 µl PBS (pH 7.0±0.2) and 20 µl 14.8 mM 4-methylumbelliferyl-ß-D-glucopyranoside (MUG) (0.3 mM in final concentration) or 10 mM 4-methylumbelliferyl-ß-D-cellobioside (MUC) (0.2 mM in final concentration) in a cuvette. The cuvette was incubated at 37 ºC for 1 h. Also, the UV spectrophotometer was maintained at 37 ºC. The absorbance (348 nm) was measured every 10 min from 0 to 60 min. Each enzyme was tested in triplicate. The extinction coefficient of MU (5370.7 $M^{-1}cm^{-1}$, at pH 7.0, temperature 25 ºC) was calculated by measuring absorbance of different standard concentrations of MU (data not shown).

2.4.6 CHU_2268 activity assay

The enzyme activity of CHU_2268 was assayed over 24 h using MUG and MUC as substrates. Lysates of CHU_2268, CenA (ß-1,4-endoglucanase from *Cellulomonas fimi*), Cex (exoglucanase from *C. fimi*) and EdinBrick I (control vector) were obtained from expression in *E. coli* JM109. A total volume of 1 ml of the mixture of lysates (5, 10, 15, 20 µl), PBS (pH 7.0±0.2) and 20 µl MUG (0.3 mM

in final concentration) [or 20 µl MUC (0.2 mM in final concentration)] was incubated at 37 ºC. The absorbance (348 nm) was measured every 1 h for 24 hours.

2.5 Endoglucanase assay (cells and lysates)

Endoglucanase activity was detected by CMC-Congo Red assay. For the lysate tests, CMC-Congo Red assay was performed by preparing PBS (pH 7.0±0.2)-Agar plates with 0.2 % CMC (Appendix I). Then, wells were drilled into the plate and 50 µl of lysate was loaded in the wells. For live cell tests, cells were streaked on 0.2 % CMC-LB-Agar with 0.38 mM IPTG and appropriate antibiotics (Appendix I). After incubation (overnight at 37 ºC), both plates were flooded with 5 ml Congo Red reagent (1 mg/ml in water). Congo Red reagent was removed after 20 min and 5 ml of 1 M sodium chloride was then poured onto the plate. After 10 min, the sodium chloride was removed and endoglucanase activity identified by the zone of clearing shown on the plate.

2.6 Exoglucanase assay and ß-glucosidase assay (cells and lysates)

For both live cells and lysates, MUC and MUG were used as the substrates for detecting exoglucanase and ß-glucosidase activities, respectively. MUC (10.0 mM) and MUG (14.8 mM) were prepared by dissolving 5 mg MUC or MUG in 1 ml of a mixture of 80 % (v/v) DSMO and 20 % (v/v) water. For live cell assays, 100 µl of 10.0 mM MUC or 14.8 mM MUG was spread evenly on LB Agar with 0.38 mM IPTG and appropriate antibiotics. Cells were streaked on the MUC or MUG plates and incubated at 37 ºC overnight. For the lysate assay, PBS (pH

7.0±0.2)-Agar plates with drilled wells were prepared with 100 µl of 10.0 mM MUC or 14.8 mM MUG. The wells were loaded with 50 µl lysate (or cell debris) and incubated overnight at 37 ºC. The plates were visualized under long wavelength UV light (365 nm) to identify the exoglucanase or ß-glucosidase activities.

2.7 Gel electrophoresis and zymogram analysis of proteins

Denaturing and non-denaturing protein gel electrophoresis was performed using the Mini-PROTEIN®II Electrophoresis Cell (from Bio-Rad). Samples for SDS-PAGE were boiled with sample buffer (Appendix I) for 4 min before loading to gels. The running buffer for SDS-PAGE (pH 8.6) was prepared as shown in Appendix I. The native PAGE gel buffer was the same buffer without SDS (Appendix I). Both SDS-PAGE and native PAGE gels (without adding SDS) were prepared at 7.5 % (w/v) acrylamide, and electrophoresis was performed at 200 V for 45 min. The native PAGE gel was in a cold condition, in which the buffer was cooled to 4 ºC and the Mini-PROTEIN®II Electrophoresis Cell was put into an ice bath while running the gel. SDS-PAGE gels were stained in Coomassie brilliant blue (Appendix I) for 40 min and destained overnight. For the zymogram experiment, the MUG or MUC were prepared at 1 mM [Kim et al. 2000] in PBS (pH 7.0±0.2), and 5 ml was sprayed on the native gel [5 cm (L) × 4 cm (W)] and incubated at 37 ºC for 1 h.

2.8 Statistical analysis and image processing

All measurement data were analyzed and graphed using MS EXCEL 2008 or SPSS Ver.16.0. TLC plates were scanned using a HP Scanjet 5590 scanner. Photographs were taken using a CANON Cybershot S90 or CANON 1Ds Mark II. In

a very few of the photographs, the contrast or exposure were adjusted using Adobe Photoshop CS4 in order to show clearer images of the outcomes.

2.9 Preliminary protein purification

Attempts were made to purify CHU_2268 from crude lysate using ion-exchange chromatography. DEAE Sepharose CL6B (Sigma) was a gift from Dr. Bruce Ward. The column packing and preparation basically followed the method of Bollag et al. (1996). Lysate (1 ml) was loaded and elution was carried out with 1 column volume (5 ml) of sodium chloride. Sodium chloride was prepared from 0 to 1 M. The flow rate was about 0.5 ml/min. Each fraction (50 µl) was mixed with 10 µl 10.0 mM MUC and 940 µl PBS (pH 7.0±0.2) incubating at 37 ºC for 30 min. The fraction with highest absorbance measurement (348 nm) was analyzed with SDS-PAGE.

2.10 Bioinformatic analysis and DNA Sequencing

Online databases and proteomics analysis tools were used to predict some characteristics of cellulases in *C. hutchinsonii*. The National Center for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/) provided the whole genome sequence of *C. hutchinsonii* ATCC 33406 and related sequence information in other species. The predicted metabolic pathways of *C. hutchinsonii* were derived from the information of the Kyoto Encyclopedia of Genes and Genomes (KEGG) website (http://www.genome.jp/kegg/). The conserved domains in cellulases and relevant domains in other species were analyzed using the BLAST service from NCBI. Other information about conserved domains was obtained from

the Pfam protein families database (http://pfam.sanger.ac.uk/). The European Molecular Biology Laboratory - European Bioinformatics Institute (EMBL-EBI) website (http://www.ebi.ac.uk/) and the Expert Protein Analysis System (ExPASy) proteomics server of the Swiss Institute of Bioinformatics (SIB) (http://www.expasy.org/) provided useful online proteomics analysis tools for predicting protein subcellular localization and amino acid sequence alignment. All online information about cellulases was obtained from three website databases: the Braunschweig Enzyme Database (BRENDA) (http://www.brenda-enzymes.org/), the International Union of Biochemistry and Molecular Biology (IUBMB) (http://www.chem.qmul.ac.uk/iubmb/) and the Carbohydrate-Active Enzymes (CAZy) database (http://www.cazy.org/). The rare codon usage calculation was performed using the service of the Laboratory for Structural Genomics and Proteomics, UCLA (http://nihserver.mbi.ucla.edu/RACC/). The signal peptide prediction was from the websites of the SignalP 4.0 server (http://www.cbs.dtu.dk/services/SignalP/), Phobius (http://phobius.sbc.su.se/) and PrediSi (http://www.predisi.de/). The restriction endonuclease cutting sites were checked by CLC Sequence Viewer Ver.6.4 or EnzymeX Ver.3.0. The designed primers were analyzed using Amplify3 Ver.3.1.4 to verify potential primer dimers or minor PCR amplification.

DNA sequencing was performed by the GenePool services of the University of Edinburgh with ABI 3730 Sanger technology platform. The results were reviewed using FinchTV Ver.1.4.0 and EnzymeX Ver3.0. The results were also examined and compared using the EMBL-EBI online multiple DNA sequence alignment tool.

Chapter 3 Cell Growth and Carbohydrate Analysis

In the experiments designed in this chapter, the growth of *C. hutchinsonii* was examined by light and electron microscopy from solid and liquid culture. Results from light microscopy suggested that *C. hutchinsonii* might attach to cellulose fibers in two distinct orientations: parallel to fibers during the early stage of degradation, and perpendicular via cell tips, in later stages. Growth of *C. hutchinsonii* was monitored with different carbon sources: glucose, cellobiose, avicel (microcrystalline cellulose) and filter paper. Glucose-based cultures gave the highest cell density. Cellobiose-based cultures caused greater acidification of the medium. Supernatant samples were assayed by thin layer chromatography (TLC). Glucose was detected in samples from the log and stationary phase of cellobiose culture suggesting that cellobiose is hydrolyzed extracellularly rather than being directly assimilated. Pentose was detected in samples from the filter paper based culture. Culture supernatants were assayed for hexose, pentose and uronic acid. The filter paper-based culture and pure filter paper showed the presence of pentose sugars, indicating that some pentose was contained in the filter paper itself. Uronic acid assays were inconclusive due to possible interference from an unidentified component; however, cellobiose-grown culture showed apparently higher uronic acid consistent with greater acidification. HPLC analysis of the filtrate from a filter paper-based culture gave inconclusive results. Further investigation is therefore required to determine the sugars released or produced during cellulose degradation.

3.1 Cell growth and microscopic examination of *C. hutchinsonii*

3.1.1 Observation of growth on plates

Cell growth of *C. hutchinsonii* was observed on DSM3T plates incubated at 30 ºC with three carbon sources: filter paper strips, 0.2 % glucose and 0.2 % cellobiose. Cell growth with filter paper was observed after 3 or 4 days as translucent yellow areas on the filter paper strips [Figure 3-1 (a) and (b)]. *C. hutchinsonii* degraded filter paper completely and produced translucent area with yellowish slime. Cells did not transfer to adjacent filter paper strips which suggested that they were unable to cross intervening agar surfaces. Degradation could be initiated either from the edge or the interior surface of filter paper strips. The plates could be maintained by putting new filter paper strips on the yellow areas.

Growth on glucose and cellobiose plates was observed after 5 to 7 days as bright yellow irregular-rounded colonies [Figure 3-1(c)]. Cells grew continuously along the streaks and were able to spread over the plate. Hence, colony counting was not possible. The ability of cells to spread over agar plates supplemented with glucose or cellobiose, but not between filter paper strips on plates lacking such supplementation, suggests that gliding is associated with substrate uptake or requires constant energy input.

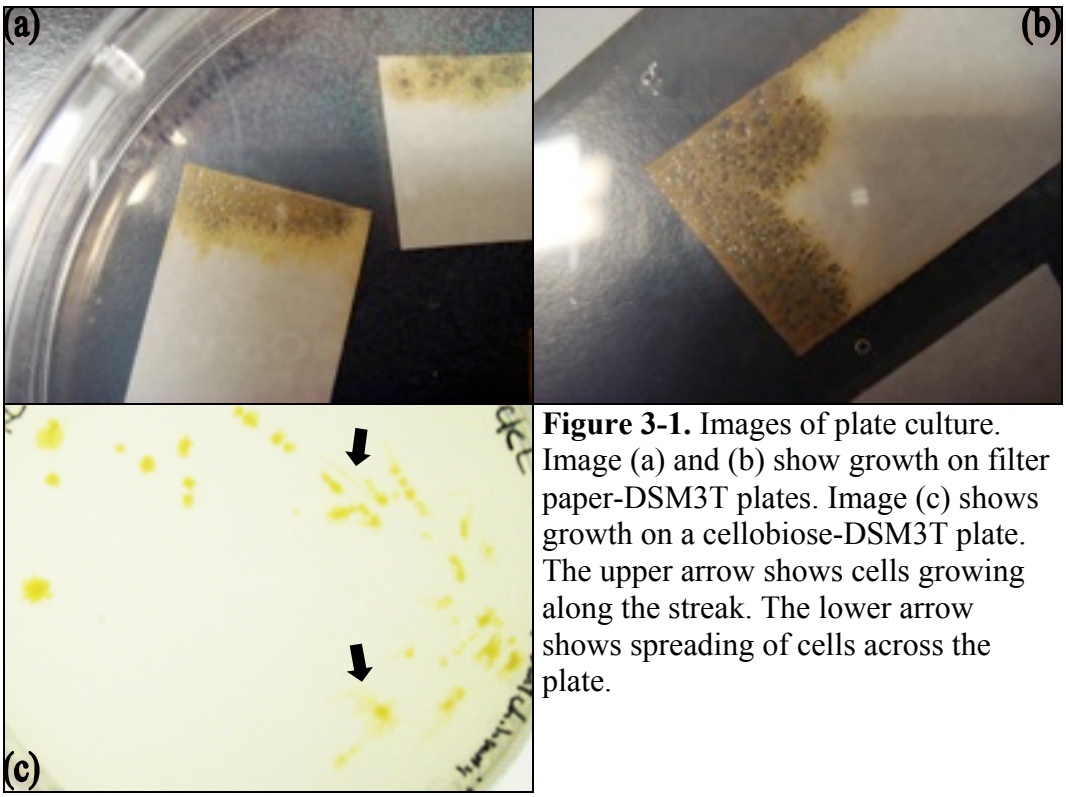


Figure 3-1. Images of plate culture. Image (a) and (b) show growth on filter paper-DSM3T plates. Image (c) shows growth on a cellobiose-DSM3T plate. The upper arrow shows cells growing along the streak. The lower arrow shows spreading of cells across the plate.

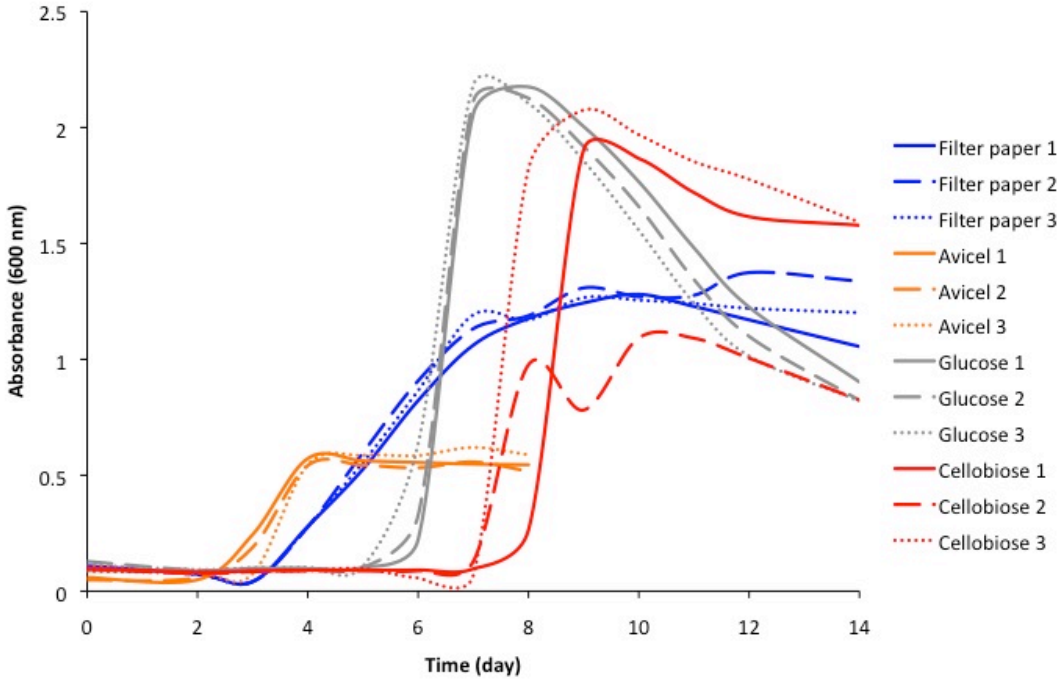


Figure 3-2. Cell growth with four carbon sources. The procedures were described in Section 2.1.1. Curves represent glucose (grey), cellobiose (red), avicel (orange) and filter paper (blue) cultures. Each carbon source was cultured in triplicate as the solid, dashed and dotted line. Data was recorded daily except for day 1 and 13.

3.1.2 Growth in liquid cultures

Cells were cultured in 100 ml DSM3T medium supplemented with 0.5 g filter paper, 0.5 g glucose, 1.0 g cellobiose or 0.2 g avicel. Each culture was prepared in triplicate. All cultures were incubated at 28 ºC, 100 rpm. Optical density (600 nm) was measured (Figure 3-2).

Lag phase on filter paper was about 3 days, on glucose about 6 days and on cellobiose about 7 days. The lag phase on avicel was about 3 days, similar to that on filter paper. The glucose cultures had the shortest log phase and declined rapidly from OD_{600} of about 2.2 to about 0.9. The maximum cell density achieved on filter paper was much lower than that seen with glucose. The growth in filter paper cultures gradually increased from day 3 and stabilized after day 9 at OD_{600} 1.1 to 1.3. The apparent lower OD and lower growth rate during log phase of filter paper culture was possibly because many cells were bound to the substrate and hence did not contribute to OD. Also, filter paper was not entirely broken down (Figure 3-3). Growth on avicel was more rapid than that seen with filter paper. Optical density of cellobiose cultures was highly variable, possibly due to flocculation, which was visibly greater in cellobiose than glucose cultures.

The pH values of all cultures were measured at the beginning and end. The pH was about 7.2 at the beginning for all four carbon sources. At the end of the experiment pH values were: filter paper 8.33 ± 0.05 (mean ± SE), glucose 7.97 ± 0.12, cellobiose 5.77 ± 0.23 and avicel 7.78 ± 0.04. Only in the case of cellobiose did pH decrease. The reason for this was unclear.

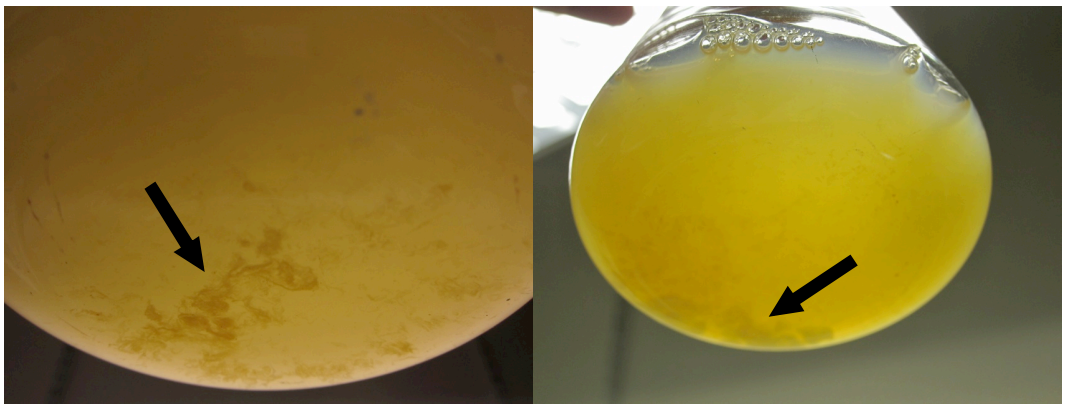


Figure 3-3. *C. hutchinsonii* liquid culture. The left image indicates the culture using cellobiose as the carbon source and the arrow shows flocs forming in the medium. The right image shows the culture using filter paper as carbon source. Particles of filter paper still remain in the medium as indicated by the arrow.

3.1.3 Microscopic examination of *C. hutchinsonii*

Results of microscopic observation from the filter paper culture are shown in Figure 3-4. Cells were attached on the surface of cellulose fibrils and were seen to attack the cellulose fibril from the terminus (white arrows) as well as the cellulose surface (black arrows).

Microscopic observation of avicel culture [Figure 3-5 (a) and (b)] showed that crystalline cellulose material was deformed badly. Clusters of cells occupied this deformed cellulose from every direction. Clusters of cells seemed to attach perpendicularly at certain sites, as indicated by the arrow in Figure 3-5 (a). Figure 3-5 (b) shows another image of an avicel particle with its crystalline texture mostly intact. At the edges of this material, cells were attacking perpendicularly, especially at the cross-section shown at the right hand arrow in this image. Parallel attachment of cells was not seen.

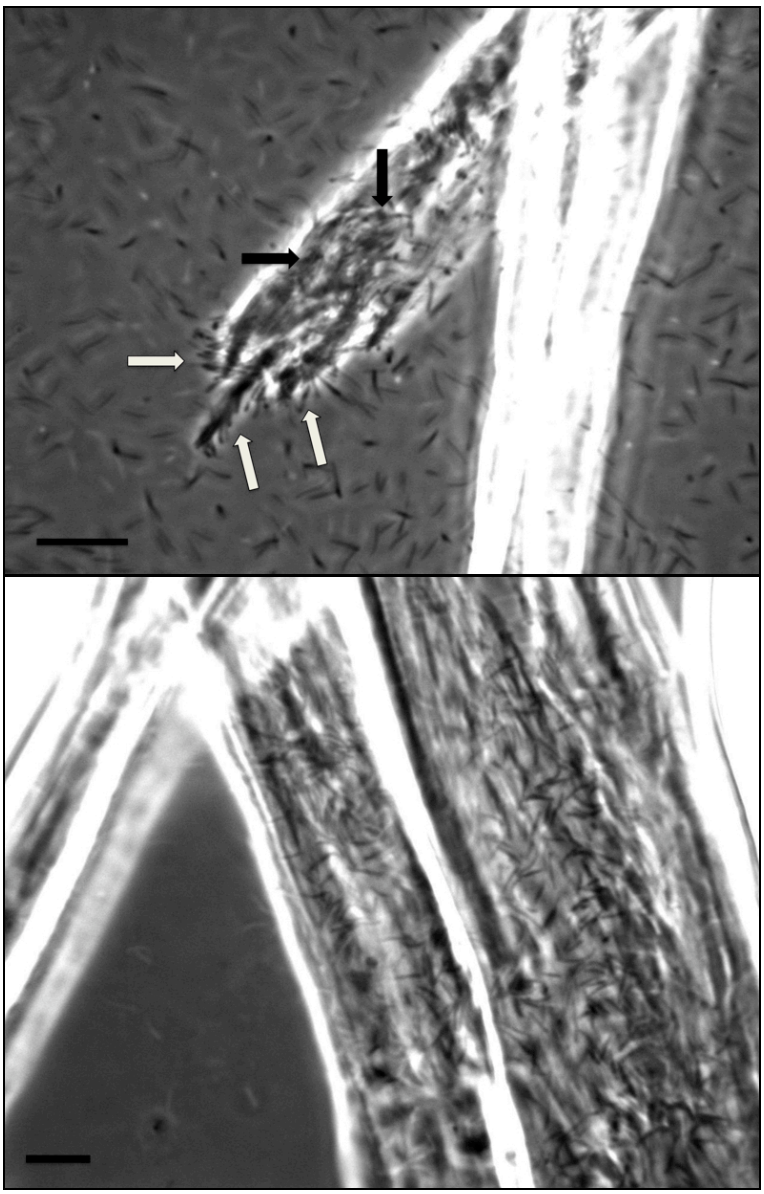


Figure 3-4. *C. hutchinsonii* images from filter paper culture. Upper image: Cells attack the terminus of a cellulose fiber where black and white arrows indicate locations of cells. Lower image: Cells attach to the surface of a cellulose fiber. Both bars in images indicate approximately 10 μm.

Thus, two different attachment modes of *C. hutchinsonii* were seen [Figure 3-5(c)]. Perpendicular attachment, with cells attached by one tip, was seen when the cellulose was deformed badly or at the cross-section of cellulose fibers, therefore might exist at the late stage of degradation. Parallel attachment was seen when the

cellulose surface was still intact, hence perhaps is associated with the early stage of degradation. The shape in this form was curved or straight and both ends of the cell were pointed. Cells with parallel attachment used their maximal cell surface to adhere to cellulose.

The yellow slime from highly degraded filter paper was examined by TEM (Figure 3-6). The image showed material branching from the cell surface and connected to rounded particles or aggregates as indicated by the arrow in Figure 3-6. This may represent EPS or other secreted materials produced by *C. hutchinsonii*.

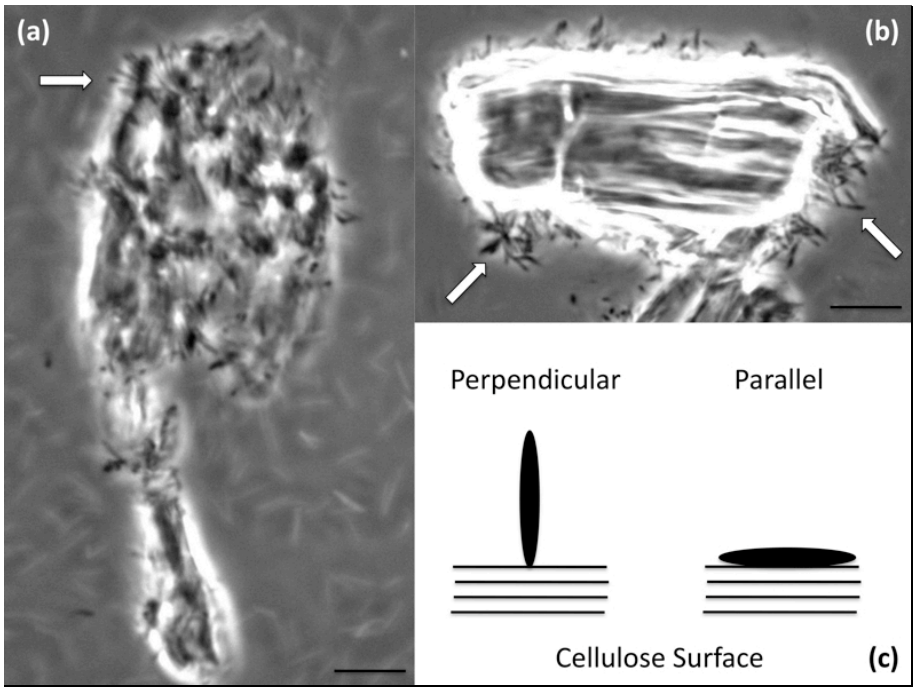


Figure 3-5. Degradation of avicel by *C. hutchinsonii*. (a). Largely de-shaped avicel particle with cells penetrating the inner level of the crystalline cellulose. The arrow indicates one site that was attacked by many cells. (b). Cylindrical avicel particle degraded by cells. The edge of the crystalline cellulose was attacked by cells. The left arrow indicates cells aggregated to one broken part of the avicel. The right arrow indicates cells attacking the cross-section of the avicel. (c). The graphic indicates two attachment modes that were used by *C. hutchinsonii*. The perpendicular style was seen at the late stage of cellulose degradation, especially when the surface of the cellulose fiber was broken through or at the cross-section of a cellulose fibril. The parallel style was seen at the early stages of degradation when the surface of the cellulose fibril was intact. Bars in image (a) and (b) indicate approximately 10 μm.

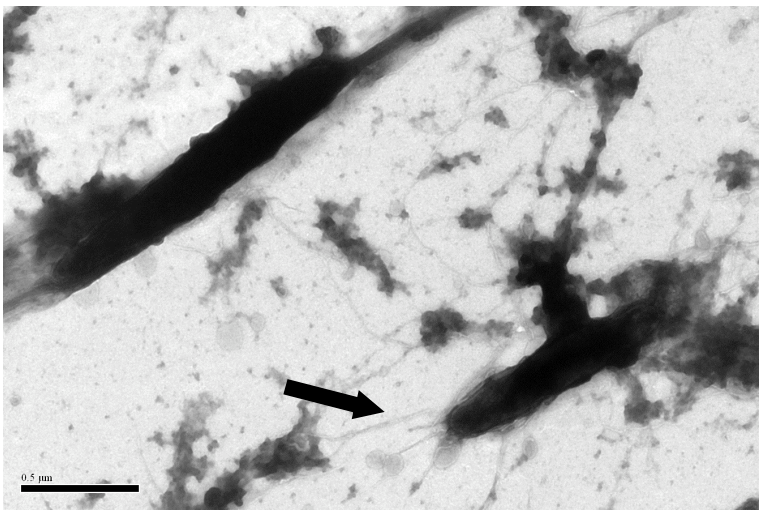


Figure 3-6. TEM image of *C. hutchinsonii*. The image was taken using cells from filter paper on an agar plate. The sample was prepared by resuspending the cells in water. The arrow indicates the secreted materials. Bar indicates 0.5 µm.

3.2 TLC analysis of carbohydrates

All TLC plates in this chapter were developed for 5 hr in water-butanol-acetic acid and sugars detected with thymol reagent as described in Section 2.1.2.3. The TLC analysis was performed with the same cultures described in Section 3.1, but each carbon source was assayed only using one repeat culture from the triplicates. The TLC results in the following sections represent the extracellular mono- and oligo-saccharides in the cultures.

3.2.1 Filter paper culture

TLC analysis of supernatants from the filter paper cultures in Section 3.1.2 is shown in Figure 3-7. The sample from day 11 for three different carbon sources was further analyzed as shown in Figure 3-8 in order to identify the sugar composition. Another sample (CH Fil) from a filter paper culture (as described in Section 2.1.1) was also analyzed by TLC following sterile filtration (0.22 µm filter) (Figure 3-9) with or without hydrolysis by hydrochloric acid (see Section 2.1.2.2).

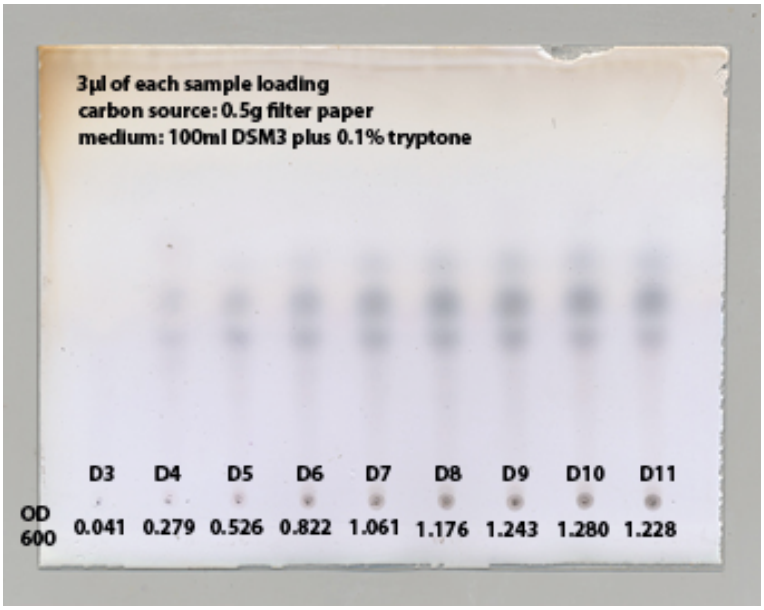


Figure 3-7. TLC of the supernatant from filter paper culture. Samples were loaded from one of the triplicate filter paper culture. Lanes from left to right are day 3 (D3) to day 11 (D11). The samples consisted of supernatant following centrifugation at 11000 *g* for 10 min. Values of culture OD_{600} are also shown.

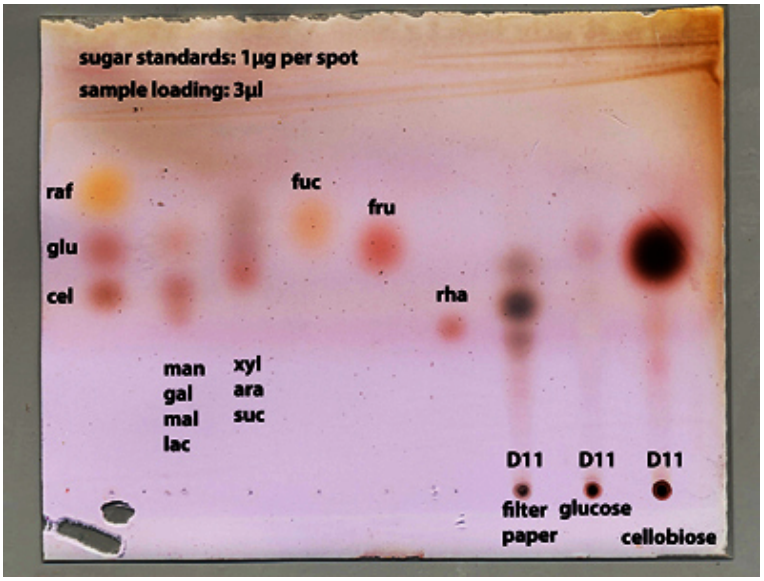


Figure 3-8. TLC analysis of supernatant from cultures grown with different carbon sources. The filter paper, glucose and cellobiose samples in this figure were the same sample from day 11 (the supernatant) as in Figure 3-7, Figure 3-10 and Figure 3-12, respectively. [Abbreviation: ara: arabinose; cel: cellobiose; fru: fructose; fuc: fucose; gal: galactose; glu: glucose; lac: lactose; mal: maltose; man: mannose; raf: raffinose; rha: rhamnose; suc: sucrose; xyl: xylose; D11: day 11.]

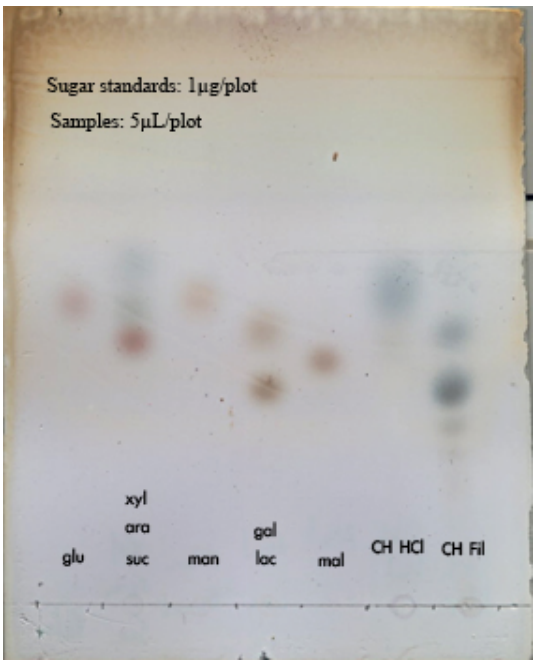


Figure 3-9. TLC analysis of filtrate from a previous filter paper culture. Eight sugar standards and two samples were included on this plate. The sample 'CH Fil' was the filtrate (0.22 µm filter) of an individual filter paper culture. The sample 'CH HCl' was the sample 'CH Fil' following hydrochloric acid digestion. [Abbreviations: ara: arabinose; gal: galactose; glu: glucose; lac: lactose; mal: maltose; man: mannose; suc: sucrose; xyl: xylose; CH HCl: sample from filter paper culture digested by hydrochloric acid; CH Fil: sample from filter paper culture with 0.22 µm filter filtration.]

Figure 3-7 shows one minor and two major spots accumulating from day 3 to day 7. All spots were bluish grey, similar to pentose standards. Figure 3-8 shows sugar standards. The spots from the filter paper culture did not exactly match any of the sugar standards; however, the color matched the color of the xylose and arabinose standards. In Figure 3-9, 'CH Fil' and 'CH HCl' were the same sample except that 'CH HCl' was digested with hydrochloric acid. Only one spot was observed from the digested sample suggesting that the other spots are oligosaccharides composed of this monomer. This single spot did not exactly match the sugar standards. However, the color of the single spot was similar to the xylose and arabinose standards suggesting this single spot might be a pentose.

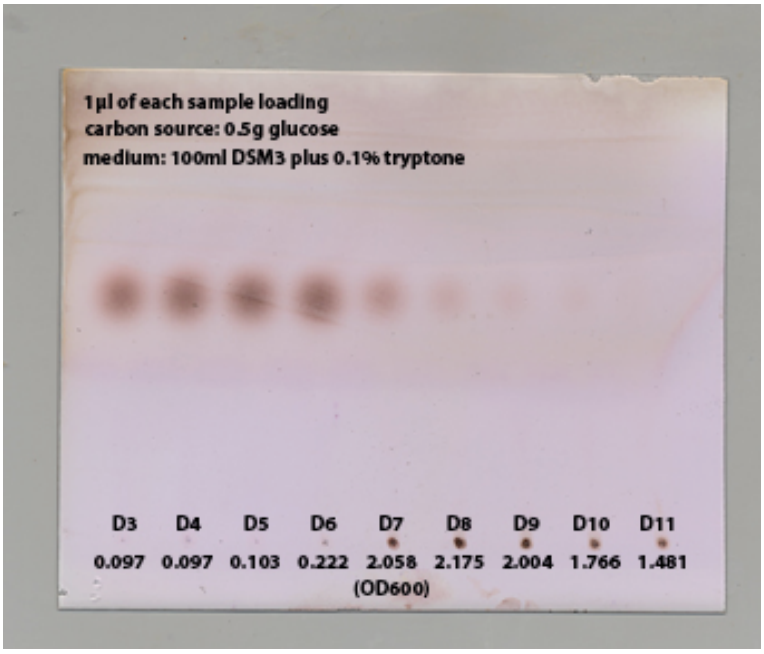


Figure 3-10. TLC of the supernatant from glucose culture. Samples were loaded from one of the triplicate glucose cultures. Lanes from left to right are day 3 (D3) to day 11 (D11). The samples were pre-treated by centrifugation at 11000 *g* for 10 min. Values of OD_{600} are also shown.

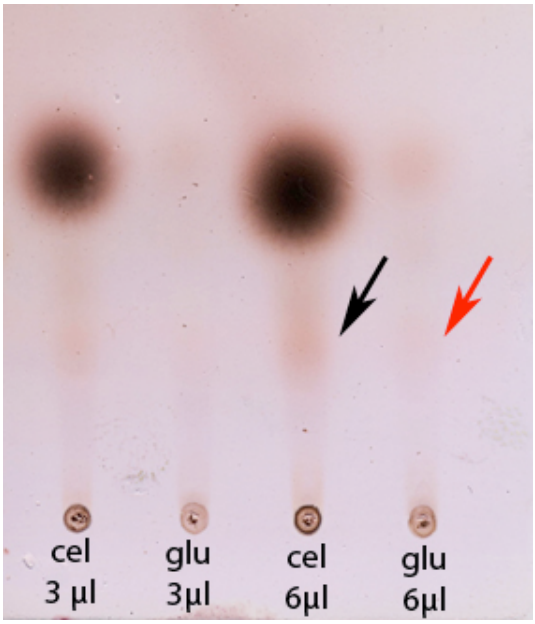


Figure 3-11. TLC plate of supernatant from glucose and cellobiose cultures (day 11). Black and red arrows indicate possible minor spots. In this TLC plate, higher volumes (3 and 6 µl) rather than 1 µl (as in Figure 3-10 and Figure 3-12) were loaded in order to visualize the possible faint spots (as indicated by the arrows). [Abbreviation: cel: cellobiose; glu: glucose]

3.2.2 TLC analysis of supernatant from glucose culture

Samples from the glucose culture of Section 3-1 were also analyzed by TLC (Figure 3-10). Only one major spot, consistent with glucose, was seen from day 3 to day 6 and this vanished after day 7, during the period where OD_{600} increased from 0.222 to 2.175, indicating that glucose was used for cell growth. No other spots were observed from this plate, indicating that no other monosaccharides or oligosaccharides were produced. Figure 3-11 shows a higher loading of the sample from day 11. A very faint glucose spot could be seen, with a possible second spot (red arrow).

Results from Figures 3-8, 3-10 and 3-11 show that *C. hutchinsonii* produced little or no mono- and oligo-saccharides during growth on glucose. EPS production in the glucose culture was not determined since TLC could not separate long chain carbohydrates.

3.2.3 TLC analysis of supernatant from cellobiose culture

A similar analysis was performed using cellobiose-grown culture (Figure 3-12). One major and one minor spot were seen between day 3 and day 7. The major spot is most likely be cellobiose, since the poor cell growth in the lag phase would lead to the carbon source (cellobiose) remaining abundant. Cellobiose was rapidly consumed by cells after day 8 and could not be detected on the plate after day 9. However, another major spot was produced while cellobiose was decreasing and its position was same as the faint spot seen before day 7. A third faint spot might be present just under the major spot from day 3 to day 8. Another TLC plate with higher

volume loading of the sample from day 11 was performed (Figure 3-11). The third spot as described above is indicated by the black arrow, but is still faint.

Figure 3-8 shows that the major spot from day 11 was consistent with glucose. Therefore cells in cellobiose-DSM3T medium might hydrolyze cellobiose first, releasing glucose, and then use glucose for growth. This suggested that cellobiose was not assimilated directly by cells. Also, the cellobiose culture showed that almost no other mono- or oligo-saccharides were produced in the culture.

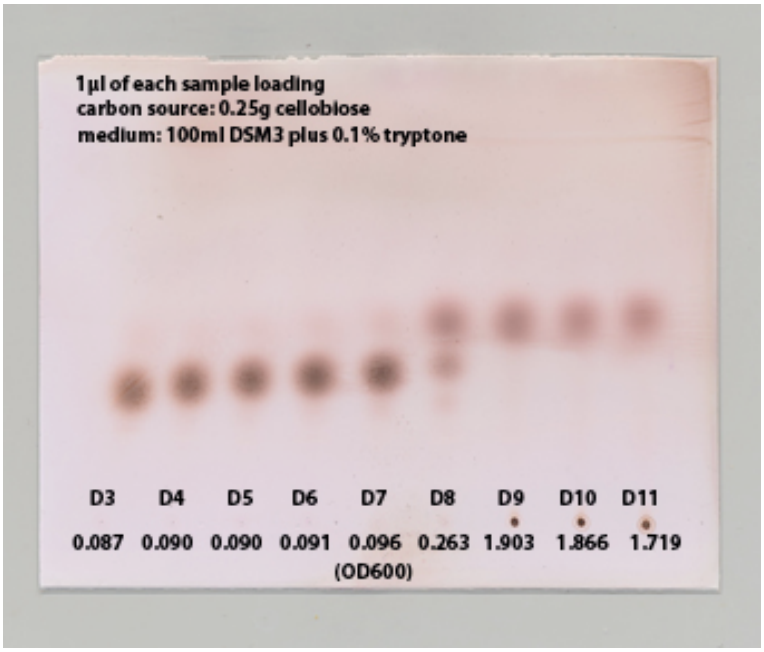


Figure 3-12. TLC of the supernatant from cellobiose culture following centrifugation at 11000 *g* for 10 min. Lanes from left to right are day 3 (D3) to day 11 (D11).

Table 3-1. Linear regression equations of colorimetric assays

Assay	Linear regression equation	R^2
Hexose	Y = 87.661 * X – 1.452	0.988
Pentose	Y = 16.454 * X – 0.186	0.996
Uronic acid	Y = 84.472 * X – 1.930	0.978

X: Absorbance measurement
Y: Units for sugar content

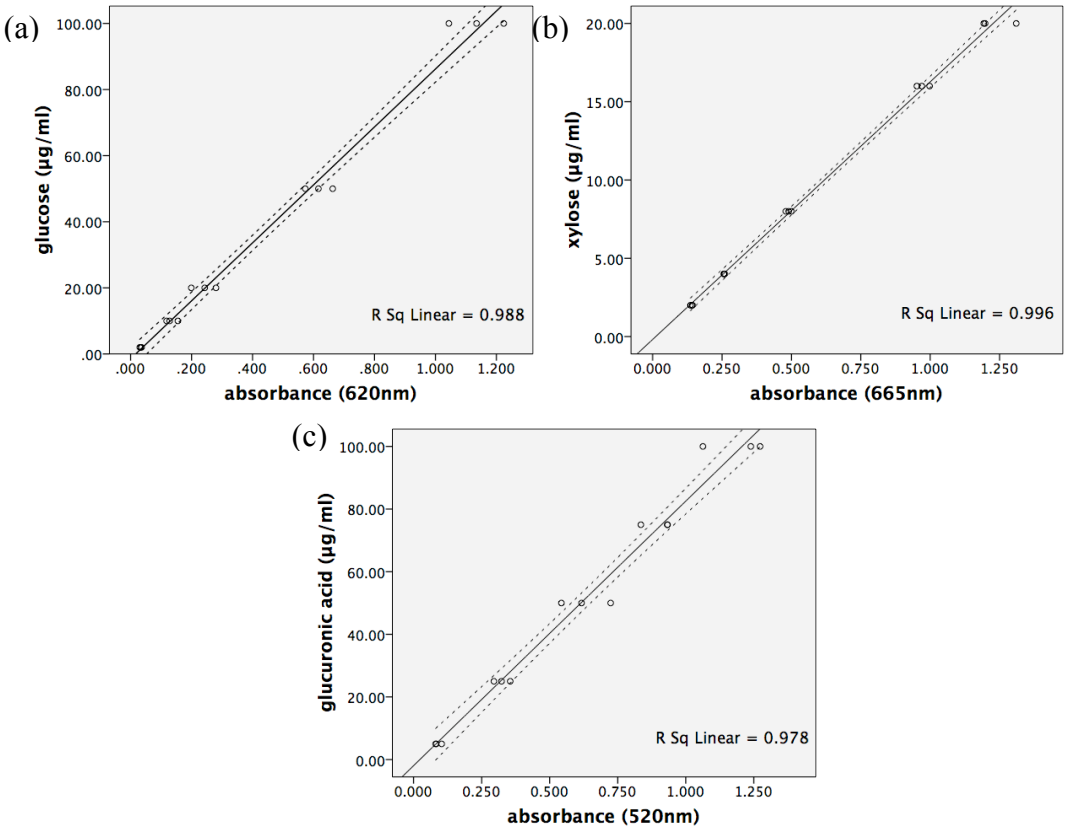


Figure 3-13. Linear regression graphs of hexose (a), pentose (b) and uronic acid (c) assay. Five concentrations of glucose, xylose and glucuronic acid were used as the sugar standards for the hexose, pentose and uronic acid assay, respectively. Every concentration of sugar standard was tested in triplicate. The dashed line represents the 95 % confidence interval of the sample means.

3.3 Culture supernatant analysis using colorimetric assays

3.3.1 Standard curve construction

Quantification of culture supernatant was performed by colorimetric assays. The anthrone-sulfuric acid method, ornicol-ferric chloride-hydrochloric acid method and borax-sulfuric acid method were applied for quantifying hexose, pentose and uronic acid, respectively. The sugar standard curves of the assays are shown in Figure 3-13, where the dashed lines show the confidence interval (95 %) of the

means. Every concentration of sugar standards was tested in triplicate. The linear regression equations and R^2 values are listed in Table 3-1.

3.3.2 Assay results

The hexose assay, pentose assay and uronic acid assay were applied to the supernatant samples from filter paper, avicel, glucose and cellobiose cultures described in Section 3.1.2. The sample of filter paper, glucose and cellobiose culture was taken from day 14, and the avicel culture sample was taken from day 8. Pure filter paper was also weighed (50 mg), shredded and digested by sulfuric acid before being tested (Section 2.1.3). The avicel culture and the pure filter paper were not tested with the uronic acid assay since not enough volume was sampled from the avicel culture and low uronic acid content was presumed in the pure filter paper. All treatments, except hexose assay of the pure filter paper, were tested in triplicate. The concentrations were calculated using the regression equations established in Section 3.3.1. Results in Table 3-2, except hexose test on the pure filter paper, were the mean values (μg/ml) of each triplicate test with ± 1 SE.

Comparing the results of hexose and pentose assay in Table 3-2, pentose content was 2x higher than hexose content in the filter paper culture, and 1.5x higher than hexose content in the glucose culture. The ratio of pentose to hexose content was 8.6 % in cellobiose culture, 77.1 % in avicel culture and surprisingly, 15.9 % in pure filter paper. However, the sum of hexose and pentose content in pure filter paper exceeded its original weight (50 mg) indicating possible cross-reaction between these two assays. In Table 3-2, the hexose content of pure filter paper was measured as 4.9 mg/ml and equaled 49 mg in the total 10 ml sample. This

measurement was close to the original filter paper input (50 mg) with about 2 % variation. Uronic acid assay tests in Table 3-2 revealed that all cultures produced small amounts of uronic acid.

Table 3-2. Colorimetric assays tests of supernatant from different carbon source culture and pure filter paper[a]

	Hexose (μg/ml)	Pentose (μg/ml)	Uronic acid (μg/ml)
Filter Paper[b]	445 ± 25	975 ± 30	35 ± 0
Glucose[b]	81 ± 20	121 ± 19	11 ± 1
Cellobiose[b]	4778 ± 596	412 ± 9	128 ± 21
Avicel[c]	53 ± 22	41 ± 2	No Data
Pure Filter Paper	4922[d]	782 ± 30	No Data

[a] : All treatments, except pure filter paper in hexose detection, were tested in triplicate.
[b] : Supernatant sampled from day 14 culture.
[c] : Supernatant sampled from day 8 culture.
[d] : Total hexose was 49.22 mg in 10 ml digested solution where 50 mg filter paper was used for acid digestion.

The color of the hexose assay (Figure 3-14), pentose assay (Figure 3-15) and uronic acid assay (Figure 3-16) with standards and samples were also examined visually. Glucose (2 to 100 μg/ml), xylose (2 to 20 μg/ml) and glucuronic acid (5 to 100 μg/ml) were used as the standards for hexose, pentose and uronic acid assays, respectively.

The results of hexose assay are shown in Figure 3-14. The filter paper culture (sample 'fp') and pure filter paper (sample 'pfp') in Figure 3-14 showed several different colors to standards, suggesting possible interference from some other component. The pentose assay is shown in Figure 3-15. Xylose standards showed a gradient of green colors. Samples of filter paper (fp) culture showed similar colors to xylose standards indicating that this supernatant contained pentose. Colors of glucose (glu) cultures and pure filter paper (pfp) were different from the glucose standard

(std glu), but matched the xylose standards indicating that the supernatant of glucose-based cultures may contain pentose, too. The pentose content of avicel and cellobiose cultures could not be identified, possibly due to the interference of glucose. The uronic acid assays are shown in Figure 3-16. The glucuronic acid showed light pink to violet pink colors; however, both filter paper culture (fp) and glucose culture (glu) showed beige colors, and cellobiose culture (cb) showed dark brown color. Hence, the uronic acid assay might not be suitable for these samples and the results obtained must be treated with caution.

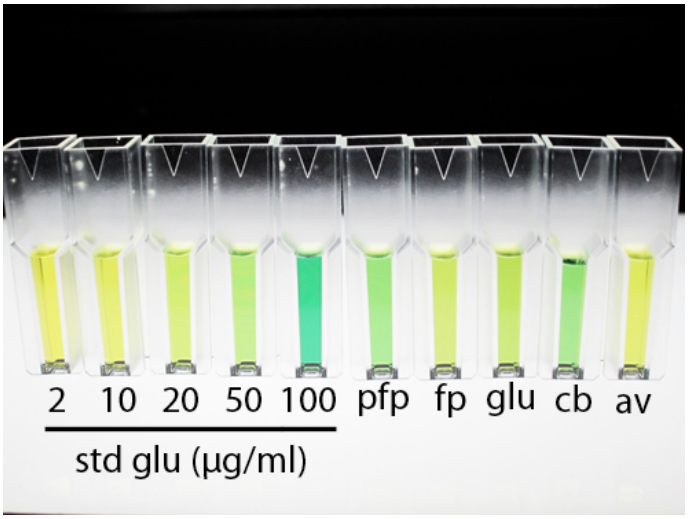


Figure 3-14. Hexose assay of sugar standards and samples. The glucose standards (std glu) were prepared in five concentrations (2 to 100 µg/ml). The supernatant samples of filter paper culture (fp), glucose culture (glu), cellobiose culture (cb) and avicel culture (av) were from the same cultures in Section 3.1.2. The sample in the pure filter paper (pfp) cuvette used sulfuric acid-digested filter paper. The 'pfp', 'fp', 'glu', 'cb' and 'av' were diluted 100x, 10x, 5x, 100x and 10x, respectively.

3.3.3 Hexose quantification of culture supernatant

The cultures described in Section 3.1.2 were analyzed using the anthrone-sulfuric acid method (hexose assay). Results are shown in Figure 3-17. The hexose content in the supernatant of the glucose and cellobiose cultures decreased. The

supernatant of the filter paper culture seemed to accumulate hexose gradually, but this might be caused by the tiny cellulose debris that could not be centrifuged.

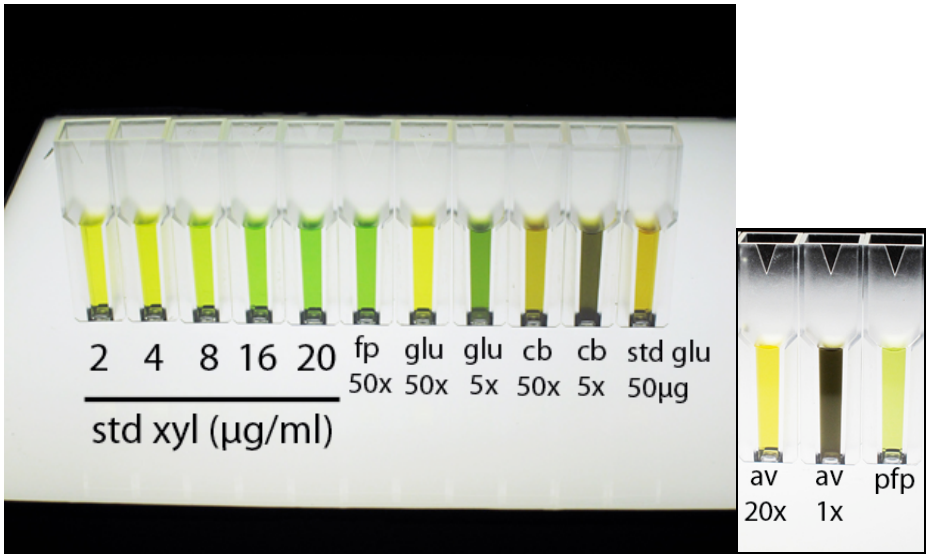


Figure 3-15. Pentose assay of sugar standards and samples. Samples of cellobiose (both cb50x and cb5x) and avicel (both av20x and av1x) showed serious interference. Glucose standard (std glu) also showed color disturbance. [Abbreviation: std xyl: xylose standards (2 to 20 µg/ml); fp 50x: supernatant from filter paper culture (50 times dilution); glu 50x and glu 5x: supernatant from glucose culture (50 and 5 times dilution); cb 50x and cb 5x: supernatant from cellobiose culture (50 and 5 times dilution); std glu: glucose standards (50 µg/ml); av 20x and 1x: supernatant from avicel culture (20 and 1 times dilution); pfp: pure filter paper digested by sulfuric acid.]

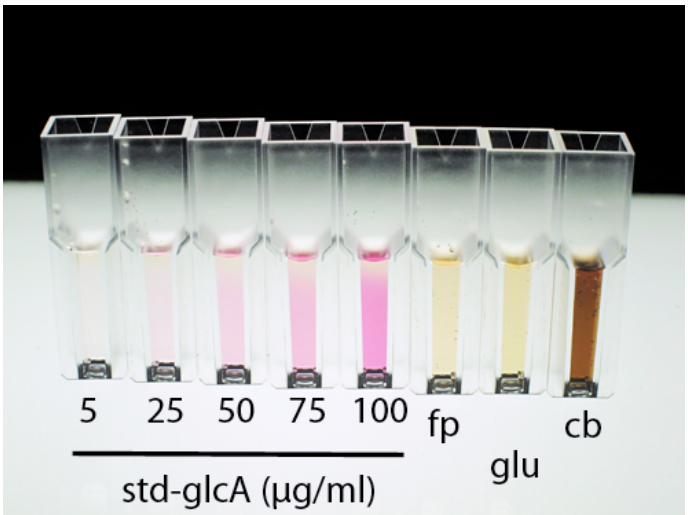


Figure 3-16. Uronic acid assay of sugar standards and samples. The cuvettes from left to right are: std-glcA: glucuronic acid standards (5 to 100 µg/ml); fp: filter paper cultures, glu: glucose culture; cb: cellobiose culture. All three samples were tested without dilution. All the supernatant samples showed color disturbance.

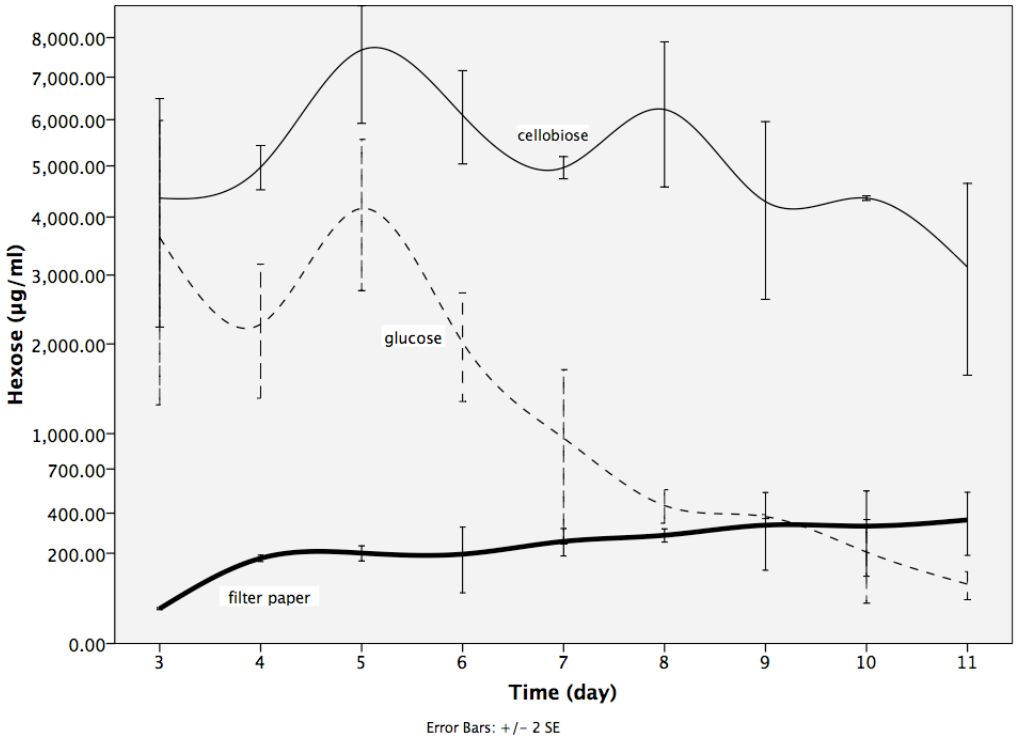


Figure 3-17. Daily changes of hexose in culture supernatants. Supernatant from the cultures of filter paper (thick-solid curve), glucose (dashed curve) and cellobiose (thin-solid curve) was tested by anthrone-sulfuric acid method. Samples were obtained from the cultures and centrifuged at 11000 *g* for 10 min. Error bars indicate ± 2 SE.

3.4 Filter paper filtrate analysis by HPLC

The filtrate from a 100 ml DSM3T filter paper (0.2 g) culture was analyzed by HPLC in two laboratories. For the Gilson HPLC in the French laboratory, the sample was filtrate from a 0.22 µm filter. One sample was prepared by further hydrochloric acid digestion as described in Section 2.1.2.2. The other sample was concentrated 10-fold by vacuum-centrifugal evaporator. Results are shown in Figure 3-18. The resolutions of the chromatograms were poor, and the noise was high. The chromatogram of the concentrated sample showed two possible peaks and the

chromatogram of HCl-digested sample showed one major peak. The two possible peaks in the chromatogram of the concentrated sample only represent the mono- or oligo-saccharides whereas the true polysaccharides would not be detected. The sugar standards were injected to verify the sugar composition in the HCl-digested sample. However, the standards were unexpectedly poorly resolved, and it was hard to determine what the sugar of the HCl-digested sample might be.

The other HPLC test was operated by Mr. Tim Gregson from Prof. Stephen Fry's laboratory. The sample pretreatment procedures were as described in Section 2.1.5.Figure 3-19 (b) showed two major unknown peaks, where one peak was eluted in a very short time, even faster than the peak of fucose as compared with Figure 3-19 (a). The other unknown peak was close to the peak of xylose in Figure 3-19 (a). There were several tiny peaks detected in Figure 3-19 (b), where five peaks could not be identified and other peaks were consistent with galactose, glucose, ribose, xylobiose and galacturonic acid. The magnified results in Figure 3-19 (c) showed other possible sugars as compared with Figure 3-19 (b). Apart from the two unknown major peaks and the peaks of galactose and glucose, two more peaks were magnified and could be rhamnose and arabinose.

As indicated in Figure 3-19, sugar composition of the filtrate from a filter paper culture could not be identified clearly. Most of the peaks were too tiny to identify and the two main peaks were unknown.

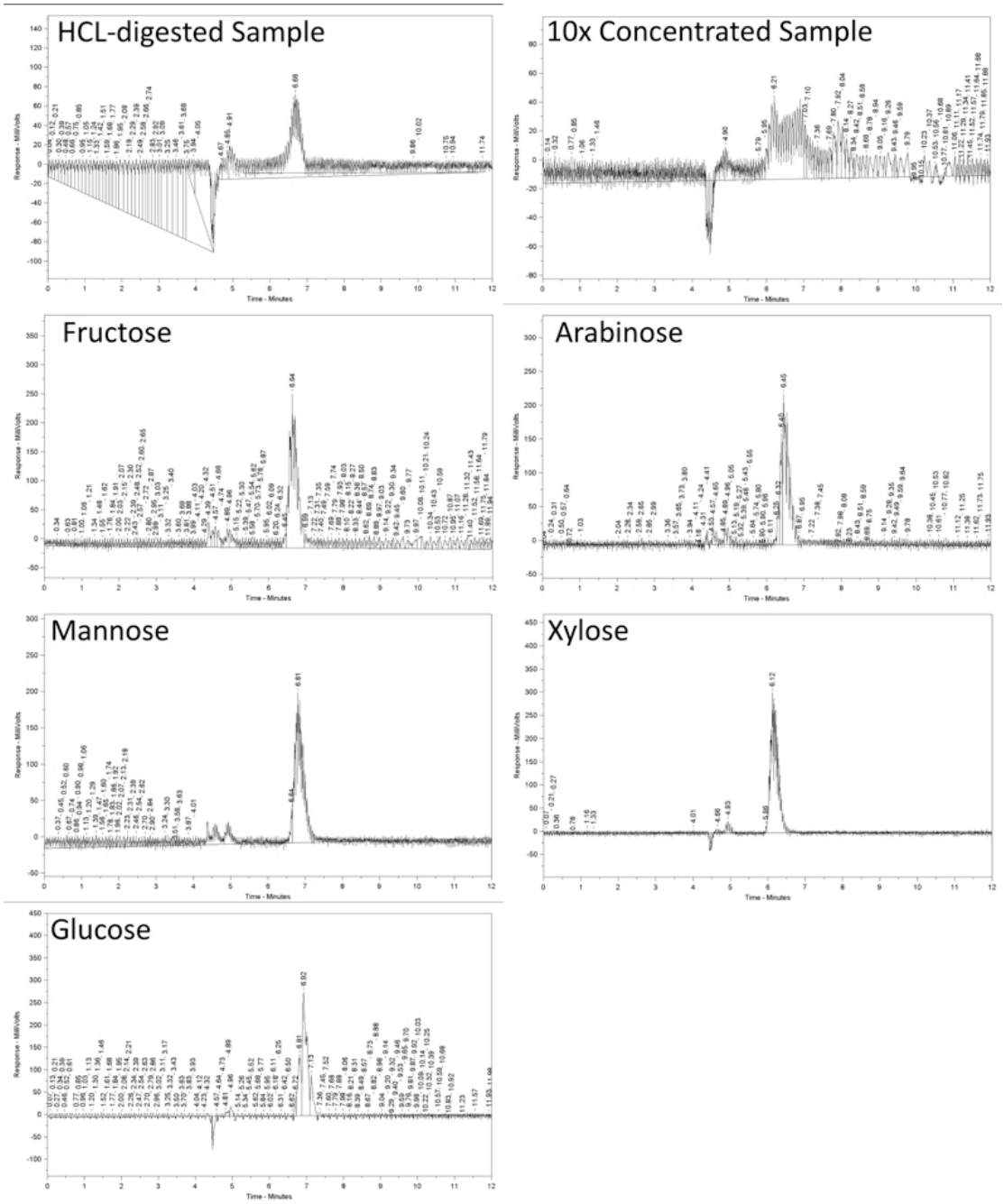


Figure 3-18. HPLC chromatograms of the filtrate from filter paper culture (French lab). The HCl-digested sample was the filtrate (0.22 µm filter) from filter paper culture with HCl digestion. The other sample from filter paper culture was 10x concentrated. Sugar standards are fructose, arabinose, mannose, xylose and glucose. The X-axis in each chromatogram represents time (min) and the Y-axis represents refractive index detector response (millivolt).

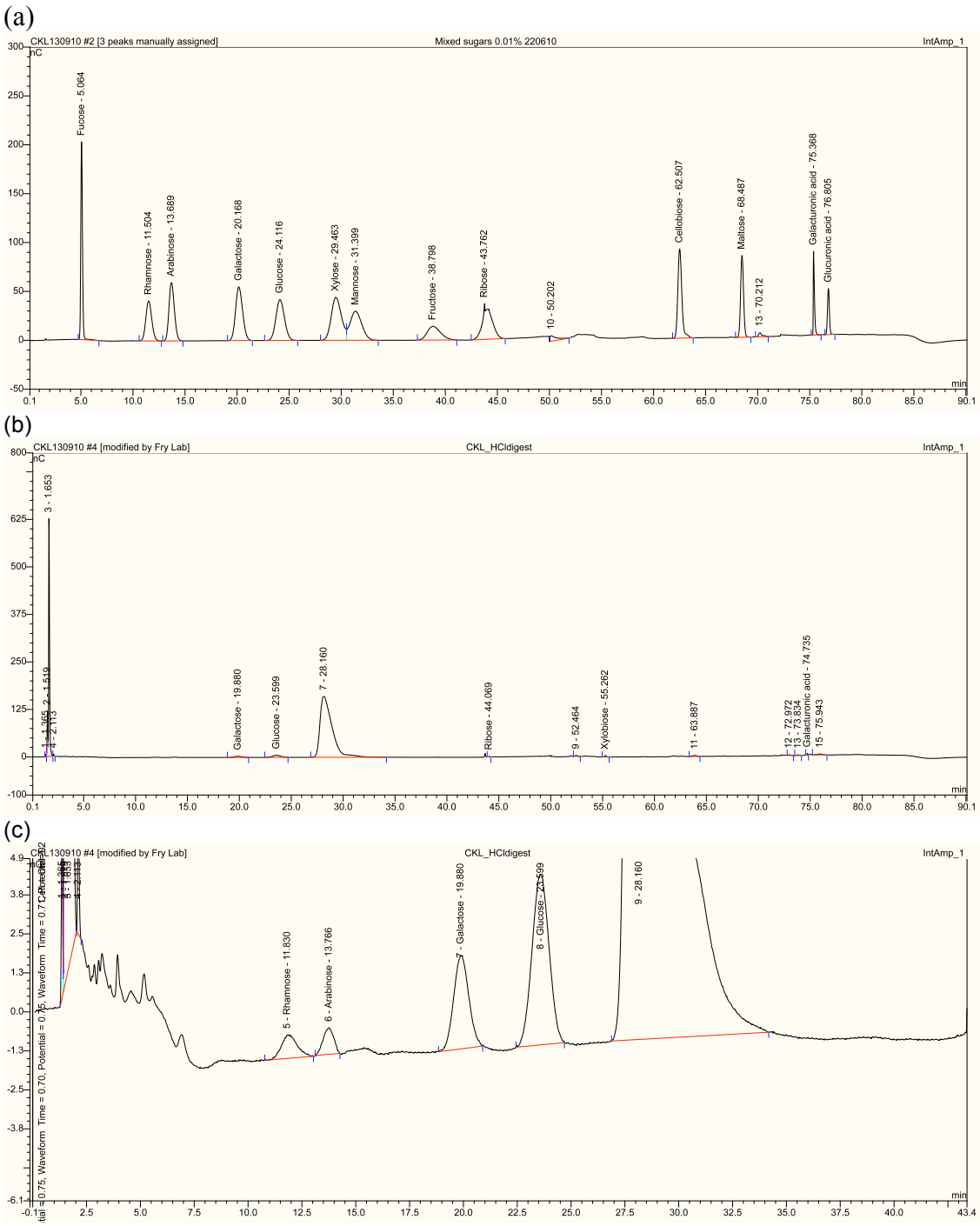


Figure 3-19. HPLC chromatograms of filtrate from filter paper culture (Fry lab). The chromatogram (a) represents the sugar standards and the sample is shown in chromatogram (b). One major peak was seen within 5 min and a minor peak was seen at 28 min. Chromatogram (c) shows the magnified chromatogram (b) in the first 43 min. The unit of X-axis was minute (min) and Y-axis was nanocoulomb (nC). The numbers shown above peaks in the chromatograms represent the retention time.

3.5 Discussion

3.5.1 Cell growth

The cell growth with filter paper, glucose and cellobiose in Section 3.1 showed different outcomes. Glucose culture showed the highest cell density (OD_{600}) measurement suggesting much of the glucose was used for cell growth and was consumed rapidly in a very short time followed by rapid decline in OD. Stationary phase was hard to identify in the glucose cultures, as in the cellobiose cultures. During log phase growth, filter paper cultures showed a lower apparent growth rate. This may have been partly due to adhesion of many cells to the filter paper strips. One of the cellobiose cultures (cellobiose 2 in Figure 3-2) showed large cell-aggregate flocs (Figure 3-3) which may have led to the low OD_{600} in Figure 3-2. Nevertheless, there were some disadvantages using optical density for cell growth measurement. For the glucose and cellobiose cultures, the measurements might be wrong due to the cell flocs and the strong yellow pigment. For the filter paper and avicel cultures, tiny substrate debris might also lead to false measurements. Therefore, measurement of protein concentration might be a better choice to for assay of cell growth in future experiments (D. K. Barnard & C. E. French, unpublished).

Glucose and cellobiose cultures in Figure 3-2 showed longer lag phase than filter paper and avicel cultures. Zhu et al. (2010) also reported that cellobiose (0.1 %) cultures (4 days) showed longer lag phase than avicel (0.1 %) cultures (1 day). Their glucose cultures (0.1 %) also grew after 1 day, which was faster than the glucose

cultures in Section 3.2.1, perhaps due to the lower glucose supply and higher rotation speed.

The pH values at the end of cell growth showed that only cellobiose cultures were acidic. The reason for this is unknown. References describing growth of *Cytophaga* spp. and other related species do not mention pH after growth. An assay showed that *Citrobacter freundii* produced strong acidification when growing on glucose, but not cellobiose (D. K. Barnard and C. E. French, not published). All cultures were examined microscopically and no contamination was observed (the microscopic appearance of *Cytophaga* is distinctive, and contamination with typical rod-shaped bacteria is easily detected microscopically). Oxidative reactions might occur in cellobiose cultures, but this could not be confirmed. However, if confirmed by further measurements, this suggests that cellobiose degradation does not follow the same pathway as glucose degradation and that cellulose has may not be degraded via cellobiose. In this regard, it is also interesting that cellobiose-grown cultures show accumulation of glucose, as well as possibly uronic acids, indicating that cellobiose is not assimilated directly.

The observation of cell growth in Figure 3-5 indicated a 'perpendicular' type of cell adhesion to substrate. Few references describe the cell growth of *C. hutchinsonii* in detail. Stanier (1942) described the degradation of cotton fibers by *C. hutchinsonii*: '…. it is possible to show that many of the attacking cells lie in these furrows, oriented in the same direction as the furrow itself.' This description and his microscopic images (in his article) were similar to Figure 3-5(a). Microscopy images from Winogradsky (1929) (written in French) only showed the young and old cultures of *C. hutchinsonii* and other related species. The shape of

Cytophaga aurantiaca shows changes described as 'black and lemon-shaped inflation' in parts of older colonies [Reichenbach 2006]. Some of the cells of *C. hutchinsonii* might similarly change shape but this needs higher resolution examination, such as TEM, to confirm.

3.5.2 Cross-reaction and limitation of colorimetric assays

Figures 3-14, 3-15 and 3-16 show possible interference in colorimetric assays for pentose and uronic acids. Pure xylose was tested in the hexose assay and showed a dark blue-green color, but the color decayed within 5 min (data not shown). Glucose produced a brown color in the pentose assay. Almost all pentose assays in Figure 3-15 showed similar color. Similar color differences were also seen in uronic acid assays. Dische (1962) reported that the anthrone method gives a positive reaction to hexoses, aldopentoses, 6-deoxyhexoses, and hexuronic acids. He also stated that the Bial reaction (pentose assay) is not specific. The 6-deoxyhexose, hexuronic acids, heptoses, trioses, and high concentrations of D-mannose and D-galactose produce green colors with absorption maxima between 650 nm and 670 nm. When D-glucose and D-fructose, which produce brown or reddish-brown colors with the orcinol reagent, are present in solution, the interference can be eliminated by dichromatic readings at 670 nm and at a second wavelength below 600 nm, at which the corresponding hexose has the same extinction coefficient as at 670 nm [Dische 1962].

Even allowing for possible interference, filter paper showed a surprisingly high content of pentose (Table 3-2, Figure 3-15). The analysis by Park et al. (2004) indicated that xylose was found in office paper. The analysis of different pulps and

papers showed less than 10 % hemicellulose, with xylose most abundant [Gutleben et al. 2004].

3.5.3 Analysis of the sugar content in cultures

Martin et al. (1968) used paper chromatography to analyze a sample from *C. hutchinsonii* culture with glucose as the sole carbon source. The result showed that the EPS of *C. hutchinsonii* contained glucose, mannose, arabinose, xylose and glucuronic acid residues. Collins (1964) cultured the anaerobic *Cytophaga fermentans* in mineral medium supplemented with 0.05 % yeast extract. His report showed that the extracellular slime produced from *C. fermentans* was composed of glucose, glucosamine, mannose and rhamnose. The EPS isolated from a *Cytophaga* sp. culture (mineral medium with 2 % glucose for 5 days) was composed of glucose (70.6 %), galactose (20.6 %), fucose (6.2 %), mannose (2.0 %), glucuronic acid (0.3 %), galacturonic acid (0.3 %) and less than 0.1 % of rhamnose and arabinose [Rättö et al. 2005]. In this chapter, three carbon sources (filter paper, cellobiose and glucose) were used for cell cultures attempting to verify the EPS composition and compare the results with Martin et al. (1968). However, results from HPLC analysis of culture supernatant hydrolysates were unclear (Section 3.4). Further attempts were made (not described in this thesis) to reduce the dense yellow pigment by dialysis (1000 MWCO), but were not successful.

Although HPLC analysis was inconclusive, TLC and colorimetric assay results still showed the variation of simple carbohydrates (mono- and oligo-saccharide) and the quantities of hexoses and pentoses in the culture supernatants.

Glucose consumption was rapid after cells started to grow (Table 3-2, Figure 3-10, Figure 3-17). Cellobiose was apparently hydrolyzed to glucose and then assimilated for cell growth (Figure 3-12). Release of glucose in cellobiose-grown cultures, but not in cultures grown on filter paper, is surprising. This suggests that cellulose assimilation may not proceed via cellobiose. No other mono- or oligo-saccharides were accumulated in glucose or cellobiose culture. Hexoses accumulated slowly in filter paper culture (Figure 3-17). These hexoses might result from either cellulose debris (disintegrated from filter paper) or polysaccharides; however, no hexose spots were observed in TLC analysis of supernatant samples (Figure 3-7). Pentoses were observed in all samples (Table 3-2), but the assay of supernatant from cellobiose culture showed interference making the result questionable (Figure 3-15). Supernatant from the filter paper culture showed a high pentose content. This may originate from the substrate itself, since about 16 % filter paper could be pentoses (Table 3-2). TLC analyses also showed the presence of oligomers of pentose in the supernatant (Figure 3-7, Figure 3-8, Figure 3-9). Although results are not entirely conclusive, consideration of TLC and HPLC results suggests that this material consists of xylooligosaccharides derived from the substrate. This is consistent with the report of Xie et al. (2007) that *C. hutchinsonii* can hydrolyze xylans but cannot assimilate xylose as a carbon source.

Chapter 4 Bioinformatics analysis of *C. hutchinsonii*

C. hutchinsonii is a cellulose–degrading bacterium; therefore its cellulase system and other related characteristics are interesting to researchers. Since its genome has been sequenced, the genetic complement of this bacterium is easily analyzed, and fortunately, many powerful software packages and online databases are available to facilitate analysis of its physiological properties. In this chapter, putative glycoside hydrolases were analyzed in order to discover candidate polysaccharide-degrading enzymes in this bacterium. More than fifty glycohydrolase-related enzymes were found, and about twenty of these were predicted to be cellulase-related enzymes, mainly ß-1,4-endoglucanases and ß-glucosidases. These putative cellulases mainly belonged to GH families 3, 5 and 9. Very few Carbohydrate-Binding Modules (CBMs) were found. Most putative cellulases only had catalytic domains without hinge linkage and CBM. Furthermore, cellulase localizations were predicted not only as extracellular but also periplasmic and outer membrane-associated. These results suggest that the general structure of cellulases and the cellulose-degrading mechanism in *C. hutchinsonii* may be different from that of other bacteria or fungi.

In addition to putative glycosylases, there are 77 predicted glycosyltransferases comprising about 2 % of the total predicted proteins in this bacterium. The possible polysaccharide transport and uptake mechanisms were analyzed as well. Some of the polysaccharide surface uptake proteins may be hidden in the hypothetical proteins.

4.1 Analysis of putative enzymes related to polysaccharide degradation

The results of genome searches for glycoside hydrolases in *C. hutchinsonii* and putative localizations of the predicted proteins are listed in Table 4-1 and Table 4-2. Classifications of GH family and Binding module (BM) in Table 4-1 and Table 4-2 were based on the NCBI and CAZy databases. The localization was determined by PSORT (Yu et al. 2010). In Table 4-1, most of the non-cellulase glycoside hydrolases were predicted to be xylanases, α-amylases or α-glycosidases. The α-amylases and α-glycosidases were predicted to be cytoplasmic, whereas the locations of xylanases were mostly not predictable using this software. The BMs in Table 4-1 belonged to families CBM4, CBM6 and CBM9. According to NCBI, CHU_0801, CHU_0959, CHU_0961, CHU_1051, CHU_1155, CHU_1238, CHU_1345 and CHU_2409 were annotated as GH family proteins as shown in Table 4-1; however, no GH domains were truly found inside these proteins through the conserved domain analysis in NCBI. Thus, this annotation was presumably based on sequence similarity to other proteins which did contain such domains. CHU_2602 was another ambiguous protein in the NCBI database. This protein was listed as both a malto-oligosyltrehalose synthase and an α-glycosidase in the same webpage. An α-amylase catalytic domain was found inside the malto-oligosyltrehalose synthase domain. The CAZy considered this protein as an α-amylase; however, KEGG still grouped this protein as malto-oligosyltrehalose synthase (EC 5.4.99.15). Also, the NCBI BLAST showed that this protein had high similarities to malto-oligosyltrehalose synthase. CHU_0981 and CHU_2802 showed similar ambiguities. These two enzymes were

both considered as ß-glucosidases belonging to GH family 16; however according to CAZy, GH family 16 does not include ß-glucosidase. The conserved domain analysis showed that one GH family 16 domain was truly present in both sequences. BLAST results did not categorize these two proteins into any type of cellulase. Hence, CHU_0981, CHU_2602 and CHU_2802 can only be considered as possible glycoside hydrolases but their true function is unclear.

According to the CAZy database, cellulases can be found in fifteen GH families, and *C. hutchinsonii* possesses representative of seven of these (GH1, 3, 5, 8, 9, 30, 74). CHU_0470 had only one conserved domain, annotated as 'GH family 3 N-terminal region' and therefore was suggested as a member of GH family 3. BLAST also showed this protein was an enzyme of EC 3.2.1.52, defined as ß-N-acetylhexosaminidase. There was only one GH family 30 (CHU_2042) and one GH family 74 (CHU_1155) hydrolase in *C. hutchinsonii*. These were both predicted to be xylanases. CHU_1240 was a putative bi-functional xylanase/esterase by protein alignment analysis and was also predicted as a GH family 8 hydrolase. CHU_3727 had a GH family 8 domain and a CBM9 domain. Protein similarity searches suggested that this protein might function as a xylanase or esterase. Both CHU_3440 and CHU_3441 had similar amino sequence length and conserved domains: GH family 8-associated at the N-terminal region and a gliding motility-associated domain at the C-terminal region. BLAST searches did not give high similarity matches; therefore the main functions of both proteins remain unclear. CHU_1075 was also a GH family 8 hydrolase with unclear function.

Table 4-1. Putative glycoside hydrolases (excluding cellulase) in *C. hutchinsonii*

Gene	Predicted Function	Localization	CD[1]	BM[2]	Por[3]
CHU_0026	Peptidoglycan-binding lytic transglycosylase	Unknown	GH23	none	No
CHU_0157	Lytic murein transglycosylase	Unknown	GH23	none	No
CHU_0353	ß-Mannanase	Unknown	GH26	CBM6	Yes
CHU_0399	Gluco-amylase-like protein	Cytoplasmic	GH15	none	No
CHU_0470	ß-N-Acetylglucosaminidase	Cytoplasmic	GH3	none	No
CHU_0801	α-Amylase	Cytoplasmic	GH57	none	No
CHU_0803	α-Glucosidase	Cytoplasmic	GH31	none	No
CHU_0959	α-Glucosidase	Cytoplasmic	GH13	none	No
CHU_0981	ß-Glucosidase	Extracellular	GH16	none	No
CHU_1075	ß-Glycosidase-like protein	Unknown	GH8	none	Yes
CHU_1155	Xyloglucanase	Extracellular or S-layer	GH74	none	Yes
CHU_1238	ß-Xylosidase/ß-1,4-endoglucanase	Unknown	Unknown	CBM9	Yes
CHU_1239	Bi-functional acetylxylan esterase/xylanase	Extracellular or S-layer	GH10	CBM4	Yes
CHU_1240	Bi-functional xylanase/esterase	Unknown or S-layer	GH8	CBM9	Yes
CHU_1345	Glycogen branching protein	Cytoplasmic	GH13	none	No
CHU_1472	Peptidoglycan hydrolase	Periplasmic	GH73	none	No
CHU_2041	Esterase/ß-xylosidase/ α-L-Arabinofuranosidase	Outer membrane	GH43	CBM9	Yes
CHU_2042	Xylanase	Unknown	GH30	CBM9	Yes
CHU_2043	Xylanase	Extracellular	GH10	CBM9	Yes
CHU_2044	ß-Xylosidase/ α-L-arabinofuranosidase-like protein	Unknown	GH43	CBM9, CBM6	Yes
CHU_2105	Xylanase	Unknown	GH10	CBM4 or 9	Yes
CHU_2290	4-α-Glucanotransferase	Cytoplasmic	GH77	none	No
CHU_2379	Xylanase	Unknown	GH11	CBM9	Yes
CHU_2409	α-Amylase	Cytoplasmic	GH13	none	No
CHU_2602	Maltooligosyl trehalose synthase	Cytoplasmic	GH13	none	No
CHU_2792	Peptidoglycan-binding lytic transglycosylase-like protein	Cytoplasmic	GH23	none	No
CHU_2802	ß-Glucosidase	Extracellular	GH16	none	No
CHU_2813	Glucosidase[4]	Cytoplasmic	GH63	none	No
CHU_3030	Peptidoglycan lytic transglycosylase-like protein	Unknown	GH23	none	No
CHU_3432	Glucosidase[4]	Cytoplasmic	GH63	none	No
CHU_3440	Glycoside hydrolase[4]	Unknown	GH8	none	No
CHU_3441	ß-Glycosidase-like protein	Unknown or S-layer	GH8	none	No
CHU_3727	ß-Glycosidase	Extracellular	GH8	CBM9	No

[1]Catalytic domain; [2]Carbohydrate-binding module; [3]*Por* secretion signal; [4]Not specified as α or ß by NCBI

There are four hydrolases in Table 4-2 for which the function cannot be predicted with any confidence. The function name of CHU_0778 from NCBI was

'glycosylhydrolase'. This sequence had an N-terminal Early set domain associated with the cellulase catalytic domain and a GH family 9 superfamily domain. The protein BLAST showed this hydrolase matching three similar ß-1,4-endoglucanases and one cellobiohydrolase (EC 3.2.1.91) from *Asticcacaulis excentricus* CB 48, the genome of which was completely sequenced by the US DOE Joint Genome Institute and submitted (unpublished) to NCBI. CHU_0961 was predicted as a GH family 9 'ß-glycosidase-like protein' and its alignment matched some hydrolases, including a ß-1,4-endoglucanase, a chitinase and a glucosamine-linked cellobiase. BLAST results indicated that CHU_2149 could be a cellulase or a xylanase. A putative GH family 5 domain was present, suggesting it should be considered as a cellulase, since GH family 5 is composed exclusives of cellulases (as indicated in CAZy). BLAST results for CHU_2852 showed a GH family 8 domain (1 ~ 400 a.a.) similar to ß-1,4-endoglucanases in some species. Although NCBI showed no conserved domains between residues 400 and 900, protein alignment results showed this section matched the chitinase domain in some enzymes.

Apart from the hydrolases mentioned in the previous paragraph, the remaining putative cellulases were predicted by NCBI to be ß-1,4-endoglucanases or ß-glucosidases (Table 4-2). Similar results were obtained from the CAZy database, but there were slight differences in the KEGG database as shown in Table 4-3. According to KEGG, the ß-1,4-endoglucanases CHU_1335, CHU_1336 and CHU_2149 were suggested as un-determined hydrolase, where CHU_1727 was considered as a licheninase rather than a ß-1,4-endoglucanase. CHU_1401 was predicted as an M42 glutamyl aminopeptidase by NCBI, but was considered as a ß–1,4–endoglucanase by KEGG. Furthermore, CHU_2802 was defined confusingly

as a ß-glucosidase or a ß-1,4-endoglucanase, although this protein might not be a cellulase as described in the previous paragraph.

Table 4-2. Putative glycoside hydrolases (cellulases) in *C. hutchinsonii*

Gene	Predicted Function	Localization	CD*	BM**	Por†
CHU_0013	ß-Glucosidase	Cytoplasmic membrane	GH3	none	No
CHU_0778	Glycosylhydrolase	Extracellular	GH9	none	Yes
CHU_0961	ß-Glycosidase-like protein	Extracellular	GH9	none	Yes
CHU_1051	ß-1,4-Endoglucanase-like protein	Unknown	Unknown	CBM6	Yes
CHU_1107	ß-1,4-Endoglucanase	Extracellular	GH5	none	Yes
CHU_1280	ß-1,4-Endoglucanase	Unknown	GH9	none	No
CHU_1335	ß-1,4-Endoglucanase-like protein	Unknown	GH9	none	Yes
CHU_1336	ß-1,4-Endoglucanase-like protein	Extracellular	GH9	none	Yes
CHU_1655	ß-1,4-Endoglucanase	Unknown	GH9	none	Yes
CHU_1727	ß-1,4-Endoglucanase	Unknown	GH5	CBM11	No
CHU_1842	ß-1,4-Endoglucanase	Unknown	GH5	none	No
CHU_2103	ß-1,4-Endoglucanase	Extracellular	GH5	none	No
CHU_2149	Glycoside hydrolase	Extracellular or S-layer	GH5	none	Yes
CHU_2235	ß-1,4-Endoglucanase	Unknown	GH9	none	No
CHU_2268	ß-Glucosidase	Periplasmic	GH3	none	No
CHU_2273	ß-Glucosidase	Periplasmic	GH3	none	No
CHU_2852	ß-Glycosidase-like protein	Unknown	GH8	none	Yes
CHU_3577	ß-Glucosidase	Unknown	GH3	none	No
CHU_3784	ß-Glucosidase	Periplasmic	GH3	none	No
CHU_3811	ß-Glucosidase	Cytoplasmic	GH1	none	No

[1]Catalytic domain; [2]Carbohydrate-binding module; [3]*Por* secretion signal

Localizations of putative cellulases were variously distributed as cytoplasmic, cytoplasmic membrane, periplasmic, S-layer or extracellular. In some cases, localization could not be predicted. Generally speaking, ß-1,4-endoglucanases were predicted mostly connecting to the outside environment and ß-glucosidases were predicted to reside inside the outer membrane. Overall, localization prediction from Table 4-2 suggested that the cellulose degradation system in *C. hutchinsonii* was neither a free-cellulase system nor a cellulosome-based system. CBM analysis in Table 4-2 also indicated that this bacterium was quite unique as compared with other

cellulose-degrading microorganisms. Only CHU_1051 and CHU_1727 had conserved CBM domains, of families CBM6 and CBM11, respectively; all other putative cellulases in Table 4-2 did not contain binding modules. Thus cellulases in *C. hutchinsonii* are not like those of a typical free-cellulase system, in which cellulases mostly have both a catalytic domain and a binding module, as in *Trichoderma reesei*. It was also unlike the cellulosome-producing organisms, such as *Clostridium* spp., since no scaffoldin proteins were discovered in *C. hutchinsonii*. And yet, *C. hutchinsonii* is a well known aerobic bacterium with high cellulose degradation ability.

Table 4-3. Differences in predictions between NCBI and KEGG

Gene	NCBI Prediction	KEGG Prediction
CHU_0778	Glycosylhydrolase	No data
CHU_0981	ß-Glucosidase	No data
CHU_1051	ß-1,4-Endoglucanase-like protein	No data
CHU_1280	ß-1,4-Endoglucanase	No data
CHU_1335	ß-1,4-Endoglucanase-like protein	Un-determined hydrolase
CHU_1336	ß-1,4-Endoglucanase-like protein	Un-determined hydrolase
CHU_1401	M42 Glutamyl aminopeptidase	ß-1,4-Endoglucanase
CHU_1727	ß-1,4-Endoglucanase	Licheninase
CHU_2149	Glycoside hydrolase	Un-determined hydrolase
CHU_2235	ß-1,4-Endoglucanase	No data
CHU_2802	ß-Glucosidase	ß-1,4-endoglucanase

Three other proteins which contained CBM-like domains were found: CHU_1221, CHU_2040, CHU_2399. CHU_1221 was identified as a putative glucose/sorbosone dehydrogenase-related protein with a DOMON_like ligand-binding domain which might be related to CBM family 9 according to CAZy. CHU_2040 was tentatively identified as an esterase-like protein with a conserved

CBM family 9 domain. CHU_2399 was a hypothetical protein which had a FA58C super family domain and was predicted to include a CBM of family 62 by CAZy.

In preparation for expression experiments, sixteen cellulase-related proteins were analyzed for the presence of codons which are rare in *E. coli* as shown in Table 4-4. The ß-1,4-endoglucanase, CenA, from *Cellulomonas fimi* was used as a reference since its expression and secretion in *E. coli* was successful. The ratio of rare codons of the selected cellulases was from 1.46 % to 4.93 %. The ratio in CenA was 2.67 %. More experiments considering rare codons are discussed in Chapter 5.

Table 4-4. *E. coli* rare codon and signal peptide analysis of selected cellulases

Cellulase Name	Rare Codons[1]	Ratio[2] (%)	Signal Peptide Prediction[3]
CHU_0013	37	3.74	Yes
CHU_1107	19	1.46	Yes
CHU_1280	11	2.58	No
CHU_1335	41	2.01	Yes
CHU_1336	25	2.56	Yes
CHU_1655	34	3.98	Yes
CHU_1727	17	2.89	Yes
CHU_1842	19	3.45	Unknown
CHU_2103	7	2.02	Yes
CHU_2149	17	1.53	Yes
CHU_2235	27	4.74	No
CHU_2268	25	3.30	Unknown
CHU_2273	15	1.83	Yes
CHU_3577	37	4.93	Unknown
CHU_3784	28	3.76	Unknown
CHU_3811	18	3.90	No
cenA	12	2.67	Yes

[1]Rare codon usage was calculated from: http://nihserver.mbi.ucla.edu/RACC/
[2]Percentage of each gene's rare codon numbers in its entire amino acid sequence.
[3]Signal peptide prediction was from: SignalP 4.0, Phobius and PrediSi.

4.2 Putative glycosyltransferase related proteins in *C. hutchinsonii*

There were seventy-seven putative glycosyltransferases distributed in thirteen glycosyltransferase (GT) families (including one group of un-characterized glycosyltransferases) comprising about 2% of the total predicted proteins. Most of the glycosyltransferases belonged to GT family 2 (28 in total) and GT family 4 (26 in total). Proteins that were predicted as GT family 2 as shown in Table 4-5 were all predicted to be ß-glycosyltransferases and were mostly (23 out of 28) predicted to be located in the cytoplasm or cytoplasmic membrane. The GT family 4 proteins in Table 4-5 were all predicted to be α-glycosyltransferases, with twenty-three located in the cytoplasm and one in the cytoplasmic membrane.

Table 4-5. Putative glycosyltransferases in *C. hutchinsonii*

Gene	GT Family	Gene	GT Family	Gene	GT Family	Gene	GT Family
CHU_0012	GT4	CHU_0863	GT2	CHU_1192	GTnc[1]	CHU_2726	GT30
CHU_0061	GTnc[1]	CHU_0864	GTnc[1]	CHU_1196	GT2	CHU_2739	GT28
CHU_0063	GT4	CHU_0868	GT4	CHU_1252	GT51	CHU_2772	GT2
CHU_0080	GT9	CHU_0869	GT4	CHU_1314	GT51	CHU_2781	GT4
CHU_0121	GT2	CHU_0870	GT4	CHU_1391	GT19	CHU_2884	GT4
CHU_0167	GTnc[1]	CHU_0884	GT2	CHU_1573	GT2	CHU_2888	GTnc[1]
CHU_0176	GT51	CHU_0885	GT4	CHU_1581	GT3	CHU_2889	GT4
CHU_0201	GT4	CHU_0888	GT2	CHU_1683	GT9	CHU_2891	GT4
CHU_0308	GT35	CHU_0889	GT4	CHU_1791	GT2	CHU_2892	GT4
CHU_0400	GT20	CHU_0890	GT4	CHU_1792	GT2	CHU_2999	GT2
CHU_0604	GT4	CHU_0894	GT4	CHU_1888	GT51	CHU_3003	GT2
CHU_0605	GT2	CHU_0895	GT4	CHU_2023	GTnc[1]	CHU_3131	GT51
CHU_0607	GTnc[1]	CHU_0896	GT2	CHU_2109	GT2	CHU_3370	GT83
CHU_0802	GT4	CHU_0912	GT2	CHU_2174	GT2	CHU_3457	GT2
CHU_0851	GT4	CHU_0913	GT4	CHU_2182	GT2	CHU_3692	GT4
CHU_0852	GT2	CHU_0928	GT4	CHU_2506	GT2	CHU_3770	GT2
CHU_0853	GT2	CHU_0929	GT4	CHU_2532	GT2	CHU_3812	GTnc[1]
CHU_0857	GT4	CHU_0930	GT4	CHU_2676	GT2	CHU_3836	GT5
CHU_0858	GT2	CHU_1044	GT2	CHU_2681	GT2	CHU_3842	GT4
CHU_0860	GT2						

[1]GT family-uncharacterized glycosyltransferase

Table 4-6. Proteins related to carbohydrate transport and uptake

	Gene	Predicted Function	Conserved Domain	LC[1]
Uptake System	CHU_0546	Outer membrane receptor; RagA protein	TonB-linked outer membrane protein, SusC/RagA family	OM[4]
	CHU_0547	Hypothetical protein	Starch binding outer membrane protein SusD	UN[5]
	CHU_0553	TonB-dependent outer membrane receptor	TonB-linked outer membrane protein, SusC/RagA family	OM[4]
	CHU_0554	Hypothetical protein	Starch binding outer membrane protein SusD	UN[5]
	CHU_1641	Hypothetical protein	TonB-linked outer membrane protein, SusC/RagA family	OM[4]
Carbohydrate exporter	CHU_0325	Sugar efflux transporter	Arabinose efflux permease	CM[2]
	CHU_0845	Polysaccharide export protein	Polysaccharide biosynthesis/export protein	UN
	CHU_0848	Polysaccharide ABC transporter, ATPase component	KpsT/Wzt ABC transporter subfamily is involved in extracellular polysaccharide export	CM[2]
	CHU_0879	Polysaccharide export outer membrane protein	Polysaccharide biosynthesis/export protein	UN[5]
	CHU_2882	Polysaccharide export protein	Periplasmic protein involved in polysaccharide export	CP[3]
Major facilitator superfamily	CHU_0367	Permease	Major Facilitator Superfamily	CM[2]
	CHU_0395	Permease	Major Facilitator Superfamily	CM[2]
	CHU_0593	Permease	Major Facilitator Superfamily	CM[2]
	CHU_0773	Fucose permease	Fucose permease, major facilitator superfamily	CM[2]
	CHU_0960	Fucose permease	Fucose permease, major facilitator superfamily	CM[2]
	CHU_1656	Glucose/galactose transporter	Fucose permease, major facilitator superfamily	CM[2]
	CHU_2160	Major facilitator superfamily permease	(none)	CM[2]
	CHU_2725	Major facilitator superfamily permease	(none)	CM[2]
	CHU_3723	Major facilitator superfamily permease	(none)	CM[2]
	CHU_1068	Major facilitator transporter	(none)	CM[2]
	CHU_3127	Transport protein	Major Facilitator Superfamily	CM[2]
	CHU_3606	Fucose permease	Fucose permease, major facilitator superfamily	CM[2]
	CHU_3844	Periplasmic protein involved in polysaccharide export	Major Facilitator Superfamily	CM[2]

[1]Localization; [2]Cytoplasmic membrane; [3]Cytoplasmic; [4]Outer membrane; [5]Unknown

4.3 Putative polysaccharide transport and uptake proteins

C. hutchinsonii is Gram-negative; therefore, materials transported from the environment to the cytoplasm must cross through the outer membrane, periplasm and cytoplasmic membrane. As illustrated in the above paragraphs, the cellulose degradation system in *C. hutchinsonii* may be different from those of previously studied microorganisms; thus, carbohydrate binding and transport is interesting.

The ABC-transport proteins support major substance transport mechanisms, and many ABC-transport associated proteins were detected in *C. hutchinsonii*, but these proteins were mostly not predicted to be related to carbohydrate transport. There are also some outer membrane proteins identified as porin-related, but not assigned to carbohydrate transport, such as OmpA and OmpH. More outer membrane proteins (or lipoproteins) were found in this bacterium but were predicted as 'unknown function'.

Sometimes, polysaccharide receptors or uptake channel systems function in carbohydrate transport. Here, as shown in Table 4-6, several proteins were predicted as 'sugar and polysaccharide transport-related' or 'carbohydrate uptake-related' according to the BLAST results. Five proteins related to the starch utilization system (SUS) of *Bacteroides* were found. Three were hypothetical proteins with conserved SUS domains, and the localization analysis for two of these were predicted as unknown. The other three SUS candidates were predicted as outer membrane-associated. Major facilitator superfamily (MFS) related proteins occasionally function in carbohydrate transport. Thirteen putative MFS proteins were found in this bacterium, four of which lacked MFS conserved domains (Table 4-6). Further BLAST analysis indicated all these four proteins were highly similar to predicted

MFS protein sequences in other species. Other MFS proteins in Table 4-6 were annotated as permease, sugar permease, polysaccharide exporter and transporters. Localization prediction showed all these MFS proteins were located in the cytoplasmic membrane.

About 2 % of proteins in *C. hutchinsonii* were found to be glycosyltransferases and were mostly predicted to be located in the cytoplasm. Therefore, mechanisms for carbohydrate export should exist. The carbohydrate exporters in the genome of *C. hutchinsonii* are listed in Table 4-6. Two exporters, CHU_0845 and CHU_0879, both possess conserved polysaccharide synthesis/export domains; however, in both cases, localization could not be predicted. Localizations of other putative carbohydrate exporters were predicted to be in the cytoplasmic membrane or cytoplasm.

Table 4-7. Localization of hypothetical proteins

Localization Prediction	Numbers
Cytoplasmic	328
Cytoplasmic Membrane	329
Periplasmic	6
Outer Membrane	49
Extracellular	51
Extracellular (Multiple Locations)	11
Unknown	861
Unknown (Multiple Locations)	53

4.4 Hypothetical protein analysis

There are 1688 hypothetical proteins in the genome of *C. hutchinsonii*. These proteins were analyzed using NCBI BLAST and PSORT. The localizations of

hypothetical proteins are summarized in Table 4-7. The annotation 'Extracellular with multiple location' in Table 4-7 refers to the protein localization predicted as extracellular or outer membrane-associated (including S-layer). 'Localization unknown with multiple locations' in Table 4-7 refers to the proteins which had cross-region localizations. For about half of the hypothetical proteins, the software could not determine the localization. Almost 40 % of the hypothetical proteins were predicted as cytoplasmic or cytoplasmic membrane and less than 7 % were outer membrane-associated or extracellular. Only six hypothetical proteins were predicted as periplasmic.

About 65 % (1090 out of 1688) of the hypothetical proteins did not show any conserved domain, and 11 % (184 out of 1688) showed conserved domains as 'uncharacterized functions'. Very few hypothetical proteins with 'uncharacterized functions' were related to other enzyme families. Twenty of these proteins might be related to the TonB-dependent siderophore receptors; one protein might be related to ß-1,4-endo-xylanase. The remaining 35 % of the hypothetical proteins showed various domains with low probability *E*-values. These included secretin and TonB N terminus short domain, CBM like domain, N- and C-terminal domain of GldM and GldN (both are gliding motility-associated proteins), 7TMR-DISM extracellular 2 domain related to carbohydrate receptor, HYR domain which might be related to cell adhesion. Some of the low *E*-value hypothetical proteins were related to the topics discussed in this chapter, such as: GH family 76 domain α-1,6-mannanase, porin superfamily that might be associated with TonB protein, SecC motif (TonB-linked outer membrane protein) which was related to SusC/RagA family domain (as introduced in the former section), UDP-glucuronate decarboxylase domain,

autotransporter ß-domain related to Type V secretion, glycosyltransferase family helical bundle domain, polysaccharide deacetylase domain, ß-amylase domain, endoglucanase E-like members of the SGNH hydrolase family, xylose isomerase-like TIM barrel, domain of starch binding outer membrane protein SusD.

4.5 Discussion

4.5.1 The missing exoglucanase

Exoglucanases (EC 3.2.1.74, EC 3.2.1.91 and EC 3.2.1.176) are distributed in GH families 1, 3, 5, 6, 7, 9 and 48. No GH family 6, 7 and 48 enzymes could be found in *C. hutchinsonii* and the GH family 1, 3, 5 and 9 hydrolases do not show exoglucanase conserved domains by BLAST analysis. Xie et al. (2007) also questioned the existence of exoglucanase in *C. hutchinsonii*, where they could not find any gene predicted to encode an exoglucanase. CHU_0778 may possess a cellobiohydrolase domain by similarity analysis. However, all the cellulase candidates including CHU_0778 are defined only based on the analysis of sequence similarity. Hence, there still lies a contradiction in *C. hutchinsonii* where a microorganism can degrade cellulose without any conserved exoglucanase in its genome.

4.5.2 Localization of Predicted Cellulases

Many hydrolases, including those with unknown localization, possess a Por secretion system domain as judged from NCBI results. The Por secretion system C-terminal sorting domain is associated with outer membrane secretion especially in

Porphyromonas gingivalis, *Fibrobacter succinogenes*, *Flavobacterium johnsoniae*, *C. hutchinsonii*, *Gramella forsetii*, *Prevotella intermedia*, and *Salinibacter rubber* [Marchler-Bauer et al. 2011]. This secretion system was shown to transport gingipains in *P. gingivalis* [Sato et al. 2005] and was discovered to translocate the gliding motility protein in *F. johnsoniae* [Sato et al. 2010]. Sato et al. (2010) also indicated that genes related to PorT protein were found in many members of the large and diverse *Bacteroidetes* phylum, including gliding bacteria such as *F. johnsoniae* and *C. hutchinsonii*. However, only one protein (CHU_2709) that was related to the Por secretion system was found in *C. hutchinsonii*. Nonetheless, the location-unknown hydrolases with Por secretion sorting peptides (Table 4-1 and Table 4-2) still gave evidence that these proteins might be located in the outer membrane.

The cellulase localization in *C. hutchinsonii* can be summarized as follows: the putative ß-glucosidases in Table 4-2 are located in the periplasmic region (CHU_2268, CHU_2273, CHU_3784), cytoplasmic membrane (CHU_0013), cytoplasm (CHU_3811) or unknown (CHU_3577, this protein should be located in the cytoplasmic membrane or periplasmic region based on PSORT analysis). Most of the putative ß-1,4-endoglucanases are predicted to be extracellular or outer membrane associated (including those location-unknown candidates). Exoglucanase is not found, unless such an activity occurs in one of the other cellulase candidates. These cellulase localization predictions indicate that the cellulose fibrils may be digested not only by free extracellular cellulase but also outer membrane (or S-layer) associated cellulases and the completion of cellulose digestion may be located inside cells, especially in the cytoplasmic membrane or periplasm. There are no scaffoldin

proteins found in *C. hutchinsonii*, suggesting that no cellulosome apparatus could be formed in this bacterium. There are possibilities for *C. hutchinsonii* to possess polysaccharide intake channels or mechanisms [Wilson 2008, 2009] if exoglucanase activity is found in those predicted periplasmic ß-glucosidases.

4.5.3 CBM in glycoside hydrolases of *C. hutchinsonii*

There are four CBMs in putative glycoside hydrolases of *C. hutchinsonii* as shown in Tables 4-1 and 4-2. These four CBMs belong to families CBM 4, CBM 6, CBM 9 and CBM 11. The structures of CBM 4 and CBM 9 are closely related [Finn et al. 2010]. Members of CBM 4 can bind to xylan, ß-1,3-glucan, ß-1,3-1,4-glucan, ß-1,6-glucan and amorphous cellulose but not to crystalline cellulose [Boraston et al. 2004, McLean et al. 2002]. Members of CBM 9 are only found in xylanases [Boraston et al. 2004]. These substrates correspond to the CBM 4 or CBM 9 possessing hydrolases in Table 4-1 where many of them are putative xylanases.

CBM 6 in *Cellvibrio mixtus* shows an unusual structure with two clefts and can bind to cellobiose, ß-1,4-ß-1,3-glucan, cello-oligosaccharides, insoluble forms of cellulose, ß-1,3-glucan, and xylo-oligosaccharides [Henshaw et al. 2004, Pires et al. 2004]. CBM 6 from *C. thermocellum* particularly binds to xylan and also has two clefts [Czjzek et al. 2001]. Members of CBM 11 are found in many bacterial cellulases [Finn et al. 2010]. CBM 11 in *C. thermocellum* can bind to both ß-1,4- and ß-1,3-1,4-glucan [Carvalho et al. 2004].

Although there are binding module candidates in glycoside hydrolases, only two were found in cellulase candidates. This leads to a question of how these cellulases bind to the cellulose fibers. It seems likely that hemicellulases, such as

xylanases, are secreted into the medium to degrade the hemicellulose matrix and expose cellulose fibers. Hence, the cell-associated cellulases can digest cellulose fibers. This is consistent with the observation that *C. hutchinsonii* can degrade hemicellulose but cannot assimilate the sugar released [Xie et al. 2007]

4.5.4 Large Substance Intake Channels

According to localization prediction, *C. hutchinsonii* may possess mechanisms for taking up or transporting cellulose chains or shorter oligosaccharides through the outer membrane. Several proteins were found and introduced in Section 4.3 that might be related to substrate uptake. Some ß-barrel-type outer membrane porins of Gram-negative bacteria transport nutrients, such as nucleotides, monosaccharides, oligosaccharides, fatty acids, urea, and short chain amides [Saier Jr. 2000]. There are other research articles showing that membrane associated sugar-binding proteins help in uptake of short oligosaccharides. Sugar-binding lipoproteins in *C. thermocellum* interact with different lengths of cellodextrins (G2 to G5) and laminaribiose [Nataf et al. 2009]. *Geobacillus stearothermophilus* T-6 has an L-arabinan utilization system and its sugar-binding lipoprotein (AnbE) can interact specifically with linear and branched arabino-oligosaccharides [Shulami et al. 2011]. Also, the final degradation of arabino-oligosaccharides in *G. stearothermophilus* is likely carried out by intracellular enzymes. In *Bacillus subtilis*, maltodextrin is bound by the maltodextrin-specific ABC transporters (Mdx-related proteins) [Schönert et al. 2006]. LacS, a galactoside-pentose-hexuronide permease, not only takes up galacto-oligosaccharide but also utilizes lactose and lactitol in *Lactobacillus acidophilus* [Andersen et al. 2011].

There is another substrate uptake mechanism discovered in *Bacteroides* spp. designated SUS, with starch- and maltooligosaccharide-binding ability [Reeves et al. 1996, Reeves et al. 1997, Shipman et al. 1999, Shipman et al. 2000]. Five putative SUS-related proteins in Table 4-6 were found in *C. hutchinsonii*, related to SusC and SusD. In the SUS mechanism, SusC is a TonB-dependent channel protein located in the outer membrane [Koropatkin & Smith 2010], and SusD is responsible for binding substrate [Reeves et al. 1996]. Another putative SUS system was found during *Flavobacterium johnsoniae* genome analysis and it was suggested that SusC-like and SusD-like proteins might function in the utilization of insoluble polysaccharides, such as chitin or hemicellulose [McBride et al. 2009]. Xie *et al*. (2007) also suggest those Sus-related proteins in *C. hutchinsonii* may be involved in binding and utilization of cellulose as well.

There are four MFS-related sugar permeases listed in Table 4-6. MSF is an important and very large transporter superfamily that was considered as a candidate for transporting sugars in early research [Maiden et al. 1987]. One family in MFS is responsible for oligosaccharide transport, such as transporting raffinose and melibiose [Pao et al. 1998]. Galazka et al. (2010) expressed two MFS proteins and one intracellular ß-glucosidase in *Saccharomyces cerevisiae* from the cellodextrin transport system of fungus *Neurospora crassa*. Results showed this expression improved growth of *S. cerevisiae* on cellodextrins and therefore indicated these two MFS proteins were playing important roles in cellodextrin uptake.

Chapter 5 Characterization of The Non-GH Family 3 Cellulases

In this chapter, nine cellulase genes were cloned and successfully inserted into vectors based on BioBrick rules. The vectors were expressed in four bacterial strains: *E. coli* JM109, *E. coli* JW2120-1, *E. coli* Rosetta and *C. freundii*. These cellulase genes were predicted as seven ß-1,4-endoglucanases, one ß-glucosidase and one GH family 5 glycoside hydrolase. For *E. coli* JM109 expression, the cell growth curves with IPTG induction were measured and, in some cases, expression was toxic to host cells after induction. Similar results were also discovered while expressing the target cellulases in the other three strains. The expression products were tested by Congo Red-CMC, MUC, and MUG assay. All cellulases expressed in *E. coli* JM109 and *E. coli* Rosetta showed no activity with Congo Red-CMC and MUC assays. MUG assays in *E. coli* JM109 and *E. coli* Rosetta showed interference from background fluorescence indicating that these strains might produce background ß-glucosidase activity. To avoid this interference, *E. coli* JW2120-1, a *bglX* mutant strain from the *E. coli* KEIO collection, was further used as the host strain for MUG assay. No ß-glucosidase activities were identified from expression in *E. coli* JW2120-1. All nine cellulases were further expressed in *C. freundii* since this bacterium was considered to have potentially better protein secretion than *E. coli*. The same assays were used. No fluorescence was observed from the MUC assays. Some of the MUG assays showed unclear results. For the Congo Red-CMC tests, CHU_1280 and CHU_1842 were the only two cellulases which showed activity in *C. freundii* expression and therefore could be considered as ß-1,4-endoglucanases.

This result also suggested that expression in *C. freundii* might be better than in *E. coli*, even though inducibility of the *lac* promoter was unclear in *C. freundii*.

5.1 The cellulases for characterization

Nine putative cellulases in *C. hutchinsonii* were chosen from three GH family groups: GH families 1, 5 and 9. These cellulases were predicted as seven ß-1,4-endoglucanases, one ß-glucosidase and one unknown type of cellulase as shown in Table 5-1. Bioinformatics of these cellulases were introduced in Chapter 4.

Table 5-1. Cellulases selected for expression

Cellulase Name	Predicted Cellulase Type[1]	Predicted Location[2]	CD[3]	BD[4]
CHU_1107	ß-1,4-Endoglucanase	Extracellular	GH5	None
CHU_1280	ß-1,4-Endoglucanase	Unknown	GH9	None
CHU_1655	ß-1,4-Endoglucanase	Unknown	GH9	None
CHU_1727	ß-1,4-Endoglucanase	Unknown	GH5	CBM11
CHU_1842	ß-1,4-Endoglucanase	Unknown	GH5	None
CHU_2103	ß-1,4-Endoglucanase	Extracellular	GH5	None
CHU_2149	Glycoside hydrolase	Extracellular	GH5	None
CHU_2235	ß-1,4-Endoglucanase	Unknown	GH9	None
CHU_3811	ß-Glucosidase	Cytoplasmic	GH1	None

[1]Cellulase type is named according to NCBI
[2]PSORT online proteomics tool in ExPASy website
[3]CD = Catalytic Domain
[4]BD = Binding Domain

5.2 Cellulase characterization from *E. coli* JM109 expression

Expression toxicity was initially tested prior to the characterization assays. After the investigation, Congo Red-CMC, MUC and MUG assays were performed

with both crude lysates and cells (grown on the plates) to identify the ß-1,4-endoglucanase, exoglucanase and ß-glucosidase activities, respectively.

5.2.1 Toxicity analysis of *E. coli* JM109 expression

The expression plasmids were constructed following the BioBrick rules as introduced in Section 2.3. In some cases, transformation efficiency in *E. coli* JM109 was very low, or transformations failed when cells were plated to plates containing IPTG. Also, numbers of transformants were at least 10 times higher on the plates with 1 % glucose, which would be expected to repress the *lac* promoter, than the plates with IPTG. These results indicated that the expression of some of the enzymes might be toxic to cells.

To analyze the expression toxicity, cell growth was monitored. Results are shown in Figure 5-1.

Data showed extremely significant differences in CHU_1107 (Student's t-test, $P < 0.001$), CHU_1655 (Student's t-test, $P < 0.001$), CHU_2149 (Student's t-test, $P < 0.001$) and CHU_2103 (Student's t-test, $P < 0.001$) as the growth curves dropped down rapidly after induction [Figure 5-1 (a), (c), (f) and (g)]. The growth of CHU_2235 was significantly suppressed (Student's t-test, $P < 0.001$) by IPTG induction. However, its induction growth curve [dashed line in Figure 5-1 (h)] showed high values of SE suggesting expression was not stable. CHU_1280, CHU_1727, CHU_1842 and CHU_3811 showed similar growth curves in the presence or absence of IPTG. There were no significant differences in these growth curves; thus the IPTG induction did not affect growth in these cases. For those genes

where expression was toxic, M9-glucose medium was applied for expression to avoid a negative impact on cell growth with IPTG induction (Section 2.4.2).

5.2.2 Characterization of ß-1,4-endoglucanase activity using *E. coli* JM109 as host strain

The transformants were picked from the blue/white screen plate and incubated overnight on 0.2 % CMC/LB plates supplemented with IPTG (0.38 mM). These plates were stained by Congo Red and washed with 1 M sodium chloride. The results are shown in Figure 5-2 (a)(b)(c), in which the positive and negative controls were CenA (from *C. fimi*) and EdinBrick I vector, respectively. Only the positive control displayed CMCase activity. CHU_1842 showed possible slight CMCase activity as a tiny zone of clearing is shown in Figure 5-2 (b).

The 5 ml initial culture in LB medium was incubated overnight at 37ºC with appropriate antibiotics. Cells were grown in 50 ml LB with 1 ml initial culture at 37 ºC. Cells were induced after 2 hr with IPTG (0.38 mM) and grown for a further 6 h. Lysate was obtained by sonication. The crude lysates were loaded to the PBS/CMC plates for overnight incubation at 37 ºC. The plates were stained with Congo Red and washed with sodium chloride (1M). The results are shown in Figure 5-2 (d)(e)(f). The positive and negative controls were still CenA and EdinBrick I vector, respectively. Every lysate displayed clearance regions, but not as large as CenA's clearance region. Also, the negative control showed almost as the same sized clearance region as the tested lysates. Therefore, none of these enzymes can be confirmed to show ß-1,4-endoglucanase activity when expressed in *E. coli* JM109.

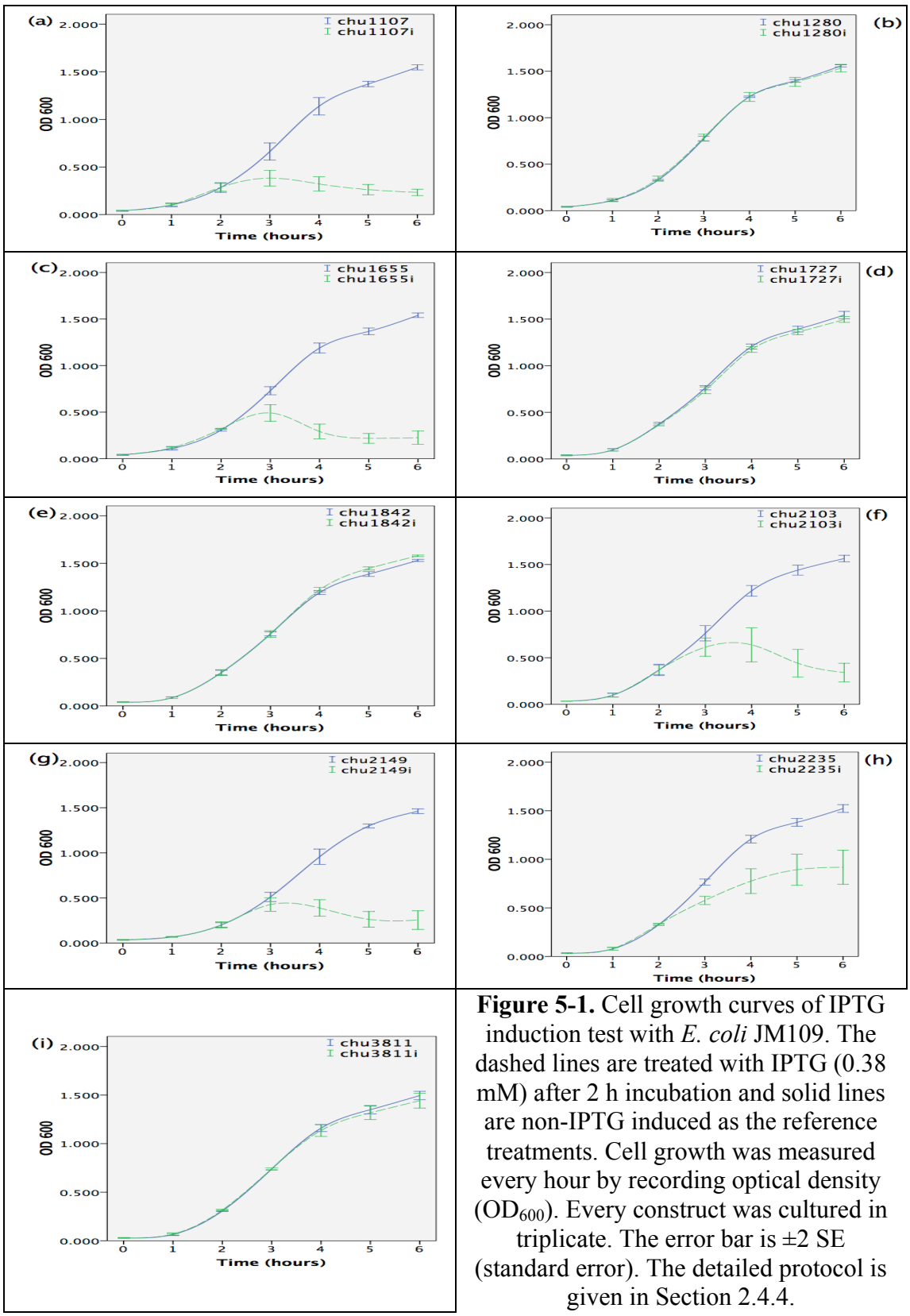


Figure 5-1. Cell growth curves of IPTG induction test with *E. coli* JM109. The dashed lines are treated with IPTG (0.38 mM) after 2 h incubation and solid lines are non-IPTG induced as the reference treatments. Cell growth was measured every hour by recording optical density (OD_{600}). Every construct was cultured in triplicate. The error bar is ±2 SE (standard error). The detailed protocol is given in Section 2.4.4.

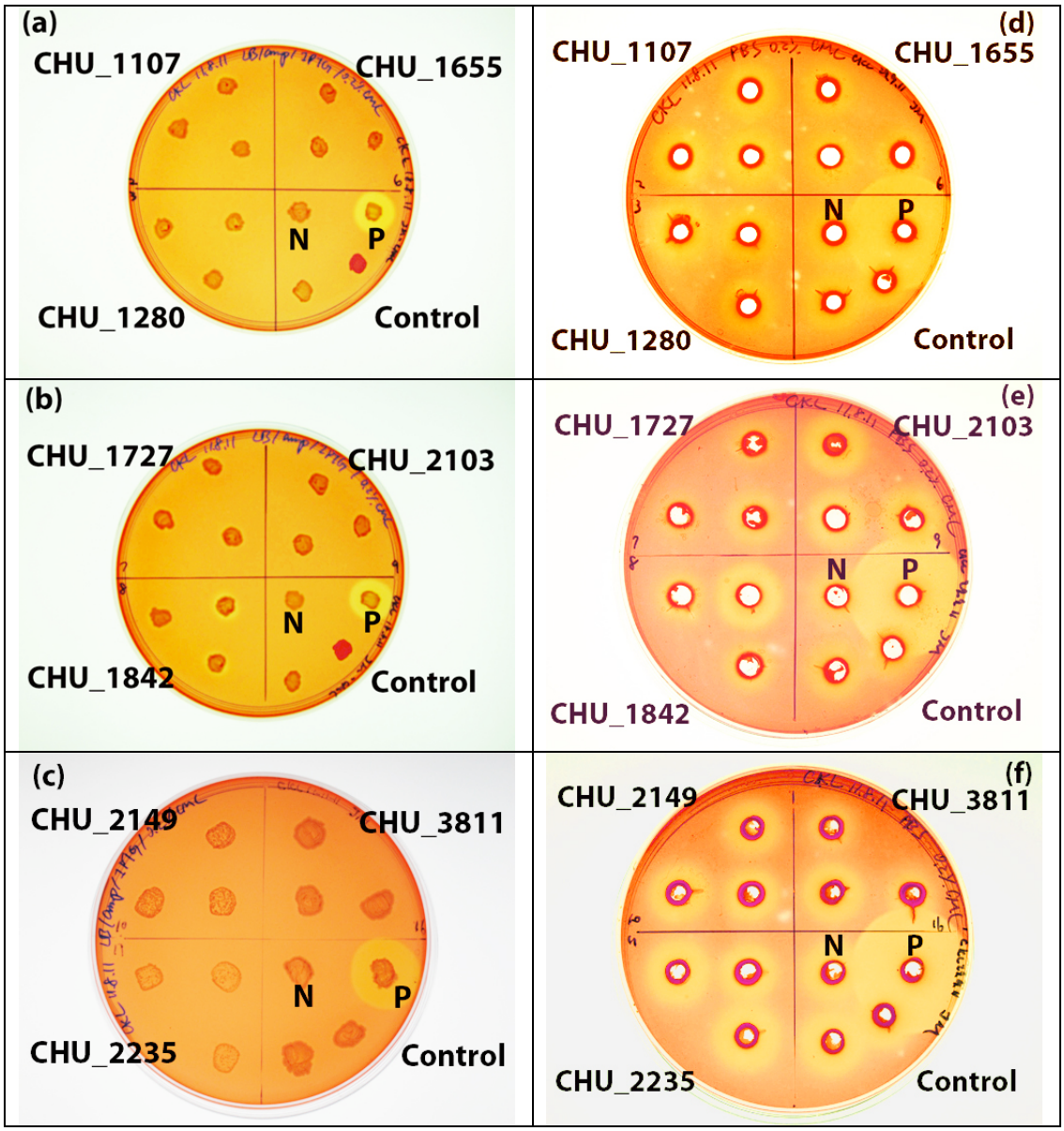


Figure 5-2. Congo Red assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* JM109. The live cells were grown on the plates with 0.2 % CMC and 0.38 mM IPTG. Lysates (50 µl) were loaded in the wells of PBS (pH 7.0±0.2)/CMC(0.2 %) plates. Each cellulase was tested (both in cells and lysate) in triplicate. The 'N' represents the negative control and 'P' represents the positive control. The remaining two wells and cell colonies in the 'Control' areas are BglX (from *E. coli* MG1655) and Cex (from *C. fimi*).

5.2.3 Characterization of exoglucanase activity using *E coli* JM109 as host strain

The exoglucanase activity was tested by MUC assay. Cells were cultured on LB/IPTG/MUC plates overnight at 37 ºC. The fluorescence (from the release of 4-MU) of the plate was detected under UV. Figure 5-3 (a)(b)(c) illustrates the MUC assay of live cells. The positive and negative controls were Cex (from *C. fimi*) and EdinBrick I vector, respectively. The results showed that none of the colonies released fluorescence (4-MU) as compared with the positive control. The crude lysates were harvested and tested using the same protocols as described in Section 5.2.2 except that CMC was replaced by MUC. Figure 5-3 (d)(e)(f) show the fluorescence from the tests of lysates using MUC as substrate. The positive and negative control were still Cex and EdinBrick I vector, respectively. No fluorescence was detected on the wells except the positive control. CenA (from *C. fimi*), as the white arrows indicated in Figure 5-3 (d)(e)(f), showed fluorescence but not as strong as Cex, suggesting that CenA may be able to degrade MUC as well. Results of MUC assays failed to demonstrate exoglucanase activity in these nine enzymes.

5.2.4 Characterization of ß-glucosidase using *E coli* JM109 as host strain

MUG was used as the substrate to test ß-glucosidase activity and the results of MUG assays are shown in Figure 5-4. The preparation of MUG assays followed the same protocols as in Section 5.2.3 except that MUC was replaced by MUG. All positive and negative controls in Figure 5-4 were BglX (from *E. coli* MG1655) and EdinBrick I vector, respectively.

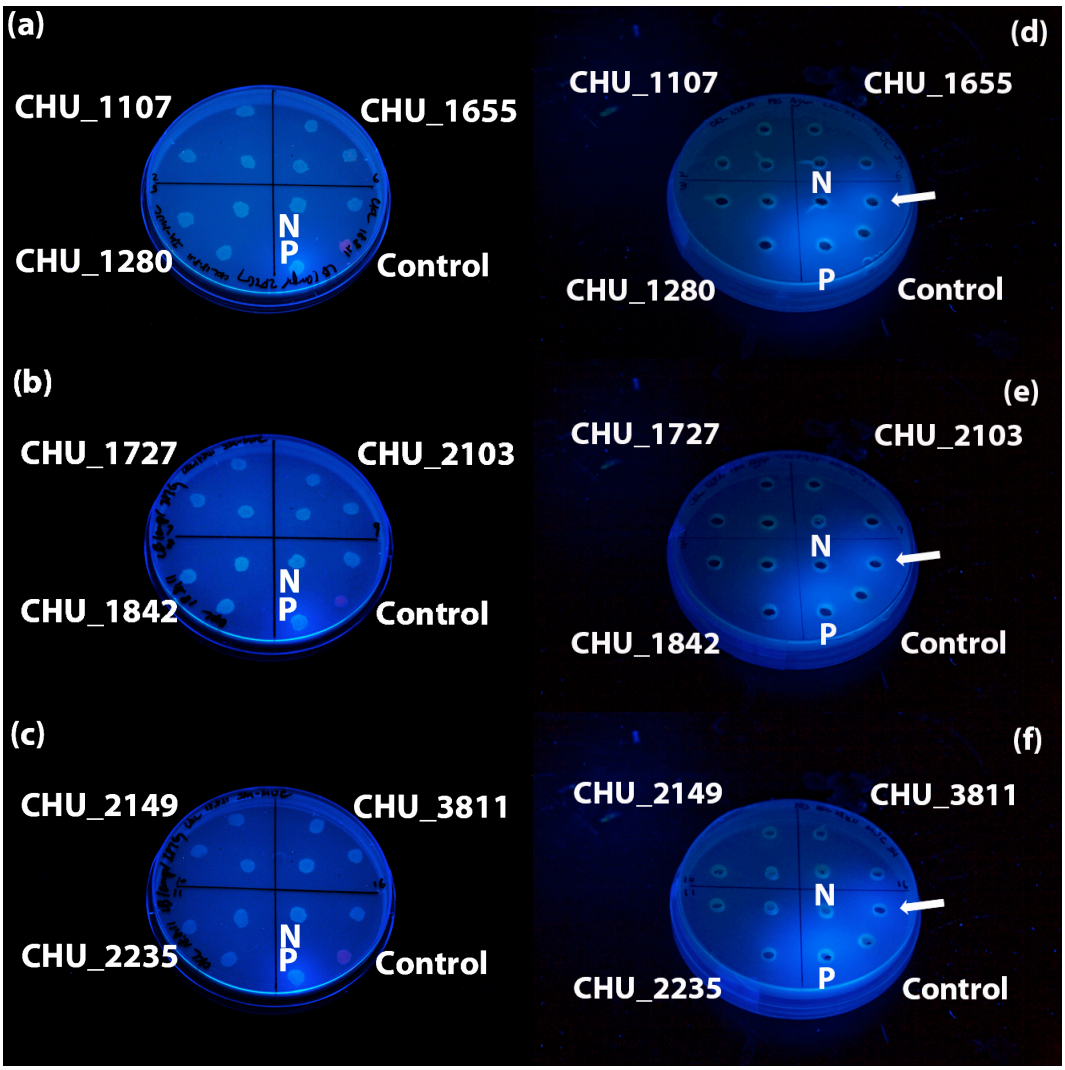


Figure 5-3. MUC assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* JM109. Each cellulase was tested (both in cells and lysate) in triplicate. 'N' represents the negative control (EdinBrick I vector) and 'P' represents the positive control (Cex, from *C. fimi*). The other two cell colonies and wells in the 'Control' areas were BglX (from *E. coli* MG1655) and CenA (from *C. fimi*). The arrows indicate the wells of CenA with minor fluorescence.

In Figure 5-4 (a)(b)(c), strong fluorescence was detected on the cells of the positive control but not on the negative control. CHU_1727 and CHU_3811 in Figure 5-4 (b) and (c) showed consistent fluorescence. Cells of CHU_1280, CHU_1842 and CHU_2103 showed inconsistent fluorescence within triplicate subcultures [Figure 5-4 (a) and (b)]. Cells of CHU_1107 and CHU_2149 showed no MUG activity as compared with the control. Cell growth of CHU_1655 and

CHU_2235 was poor, maybe due to the toxicity of IPTG induction on the plate. Fluorescence of both CHU_1655 and CHU_2235 cells was as faint as the negative control. In Figure 5-4 (d), lysate of CHU_1107 and CHU_1655 showed no fluorescence (compared to the negative control). These two lysates were both expressed in M9 medium due to the toxicity of IPTG induction in LB medium. Nevertheless, no evidence indicated that expression in M9 medium gave lower MUG background interference than expression in LB medium. In Figure 5-4 (d), (e) and (f), all of the lysates (except CHU_1107 and CHU_1655) showed partial MUG activities, and both positive and negative controls showed strong MUG activities.

Results of live cell tests and lysate tests with MUG were not consistent with each other. The negative control showed MUG activity in lysate tests but not in live cell tests. Several target genes showed fluorescence with live cell tests but not with lysate tests and vice versa. Therefore, these results could not determine whether any ß-glucosidase activity was expressed. Furthermore, results of the negative control suggested that *E. coli* JM109 might have ß-glucosidase activity expressed in its crude lysate. Thus, those MUG-positive results might contain this ß-glucosidase background interference.

5.3 Cellulase characterization from *E. coli* Rosetta expression

According to the results in Section 5.2, no obvious cellulase activities were revealed and in some cases, expression was toxic to host cells after induction. Thus, *E. coli* Rosetta strain (Novagen) was then applied for expression in order to reduce the potential negative effect that was produced by rare codon usage (as discussed in Chapter 4). This strain was also applied to improve the enzyme productivity from

expression. The *E. coli* Rosetta strain is basically an *E. coli* BL21(DE3)pLysS strain with additional supplement of tRNAs for the rare codons on a chloramphenicol-resistant plasmid, pRARE2. Nonetheless, the constructed plasmids with the *lac* promoter could still be applied in *E. coli* Rosetta.

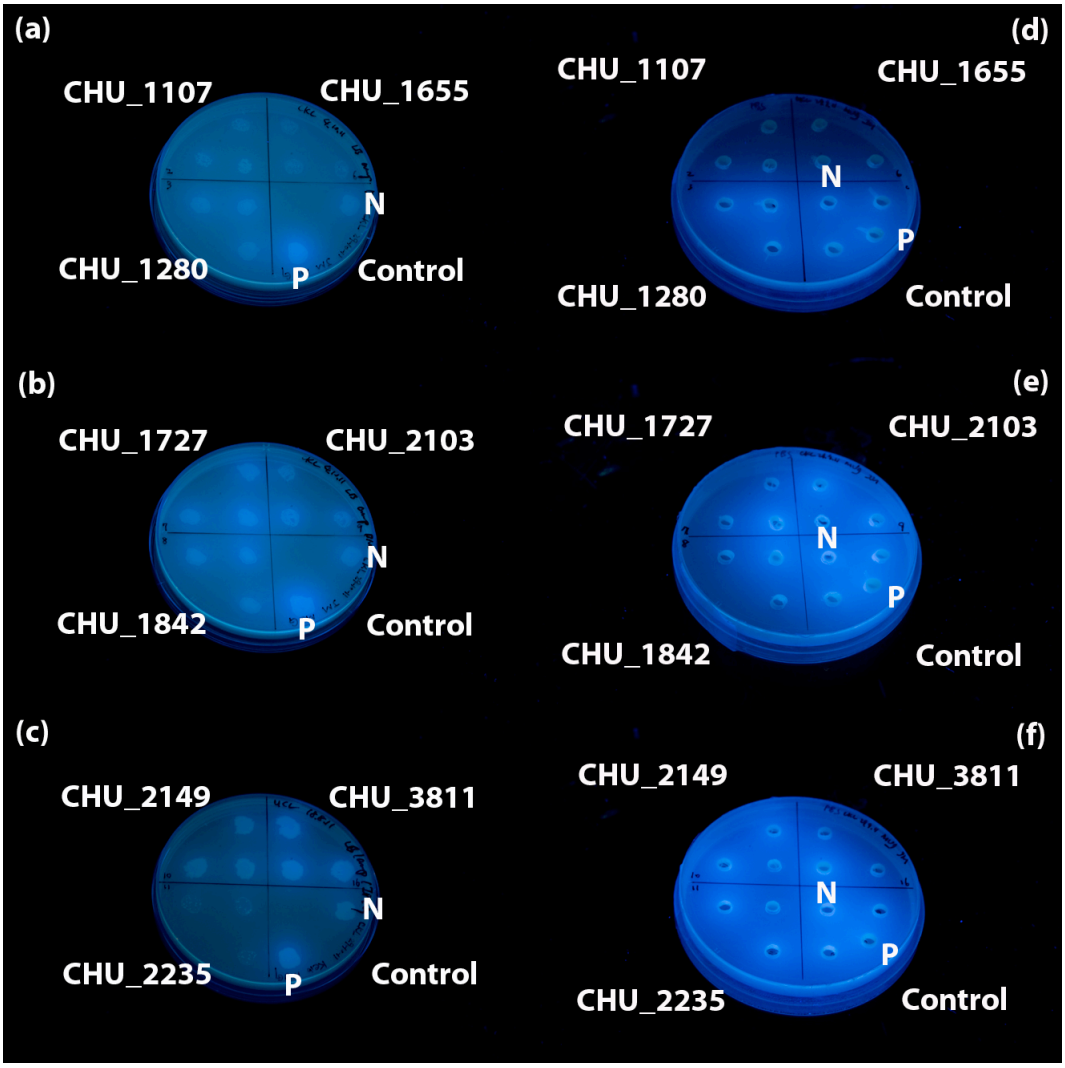


Figure 5-4. MUG assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* JM109. Each cellulase was tested (both in cells and lysate) in triplicate. 'N' represents the negative control and 'P' represents the positive control. The other two wells in the 'Control' areas in (d)(e)(f) are Cex and CenA (both are from *C. fimi*). Cell growth of CHU_1655 [as shown in (a)] and CHU_2235 [as shown in (c)] was poor due to the toxicity produced by IPTG induction.

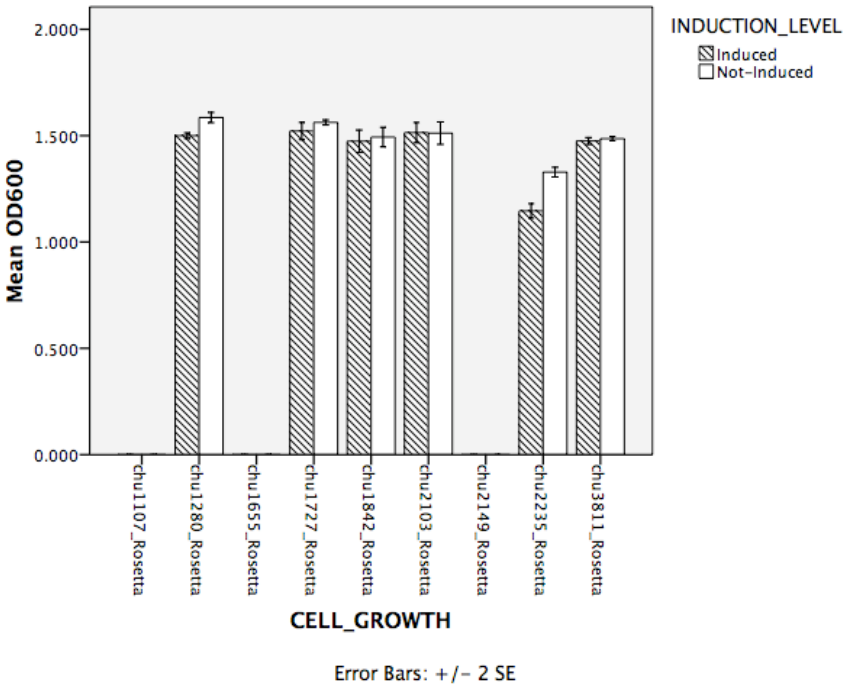


Figure 5-5. Bar chart of toxicity in *E. coli* Rosetta following IPTG induction. The plasmids of CHU_1107, CHU_1655 and CHU_2149 failed to transform *E. coli* Rosetta, therefore there were no data for the analysis in this Figure. Each culture was prepared in triplicate. The mean value of final OD_{600} was measured and calculated after 6 h expression.

5.3.1 Toxicity test of *E. coli* Rosetta expression

The same plasmids used for *E. coli* JM109 expression were introduced into *E. coli* Rosetta. Only six transformations were successful. CHU_1107, CHU_1655 and CHU_2149 failed to transform even with additional glucose (1 %), which was used to suppress minor expressing leakage of the *lac* promoter. The successful transformations were used to test the expression toxicity (protocols were given in Chapter 2) and the results are shown in Figure 5-5.

No data could be analyzed for CHU_1107, CHU_1655 and CHU_2149 due to failure of the transformation. The optical density (OD_{600}) of other cultures showed that only CHU_2235 might be inhibited after IPTG induction as shown in Figure 5-5. Compared with the five non-induced cultures in Figure 5-5, the mean optical density

of CHU_2235 was still lower than others indicating that cells were harmed even without induction.

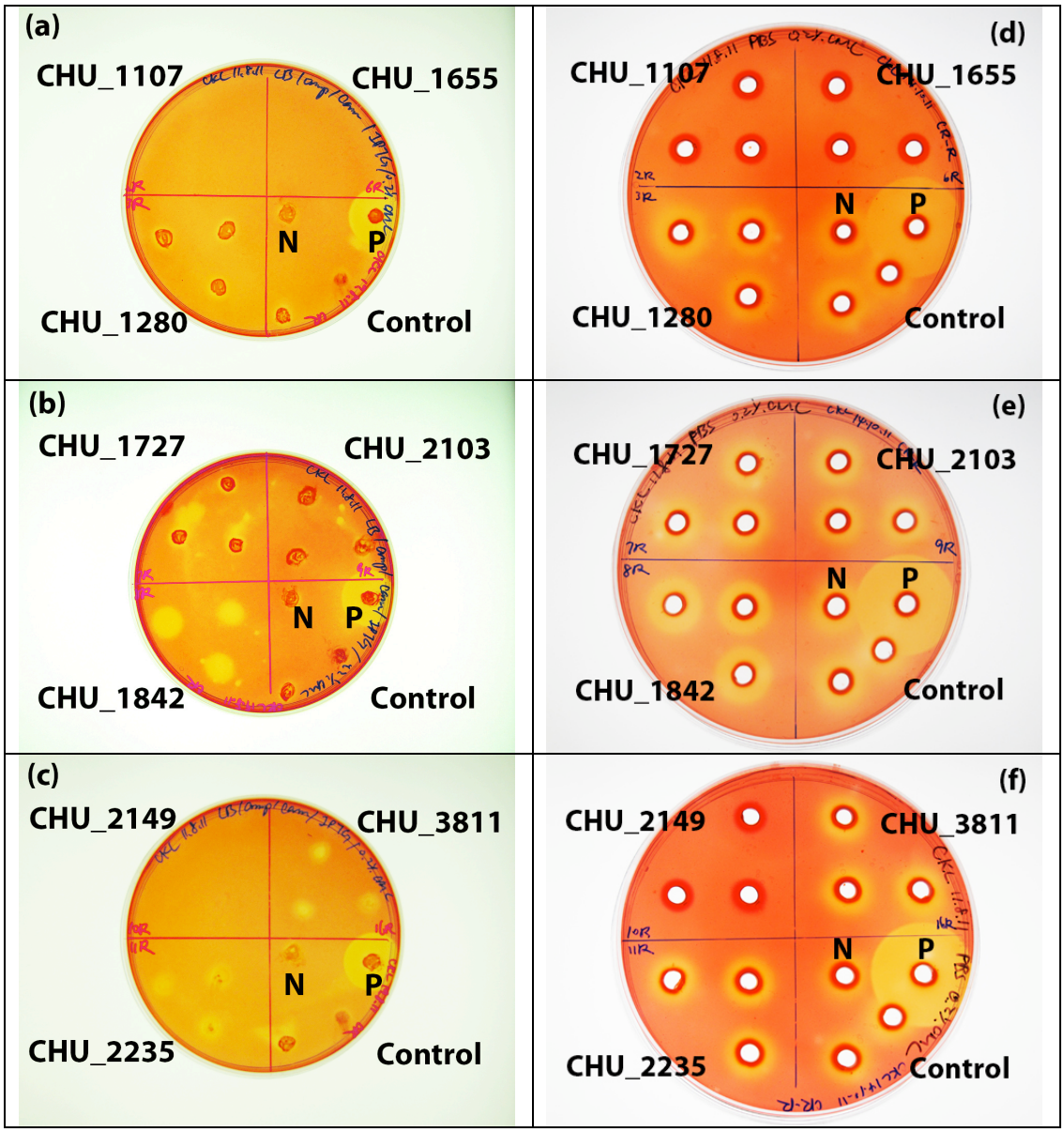


Figure 5-6. Congo Red-CMC assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* Rosetta. Both areas (live cells and lysate) of CHU_1107, CHU_1655 and CHU_2149 were empty due to the failure of transformation. In (b) and (c), cells of CHU_1842, CHU_2235 and CHU_3811 were washed away by 1 M sodium chloride. Positive controls (P) were CenA (from *C. fimi*) and negative controls (N) were EdinBrick I vector. The un-labeled cell colonies and wells in the 'Control' areas were Cex (from *C. fimi*) and BglX (from *E. coli* MG1655).

5.3.2 Characterization of ß-1,4-endoglucanase activity using *E coli* Rosetta as host strain

The protocols and controls of Congo Red-CMC assays for both treatments (lysate and live cell) were the same as in Section 5.2.2. For the live cell tests, the sodium chloride-removed plates were photographed and are shown in Figure 5-6 (a), (b) and (c). The empty areas for CHU_1107, CHU_1655 and CHU_2149 were because of the failure of transformation as described in Section 5.3.1. The poor growth of CHU_2235 in Figure 5-6 (c) suggested that this construct was still toxic to the host cells. The yellow clearance areas without cells of CHU_1842 and CHU_3811 were because the cells were washed away when immersed with 1 M sodium chloride. From the results in Figure 5-6 (a),(b) and (c), no ß-1,4-endoglucanase could be identified from the live cell tests.

As in Figure 5-6 (a)(b)(c), no lysate was tested with CHU_1107, CHU_1655 and CHU_2149 in Figure 5-6 (d)(e)(f). Among the six lysates in Figure 5-6 (d)(e)(f), only CHU_1842 showed slightly larger clearance (2 out of 3 wells) than the other five target lysate and negative control. But still, the positive control had the largest clearance area of all. The clearance rings of the negative control suggested that *E. coli* Rosetta and *E. coli* JM109 showed the same background in the Congo Red-CMC assay.

5.3.3 Characterization of exoglucanase using *E coli* Rosetta as host strain

Exoglucanase activity was determined by MUC assay. The protocols and controls were the same as in Section 5.2.3. Again, no lysate or cell culture could be

harvested for MUC assays of CHU_1107, CHU_1655 and CHU_2149 due to the transformation failure. The MUC assays with live cells and lysate are shown in Figure 5-7.

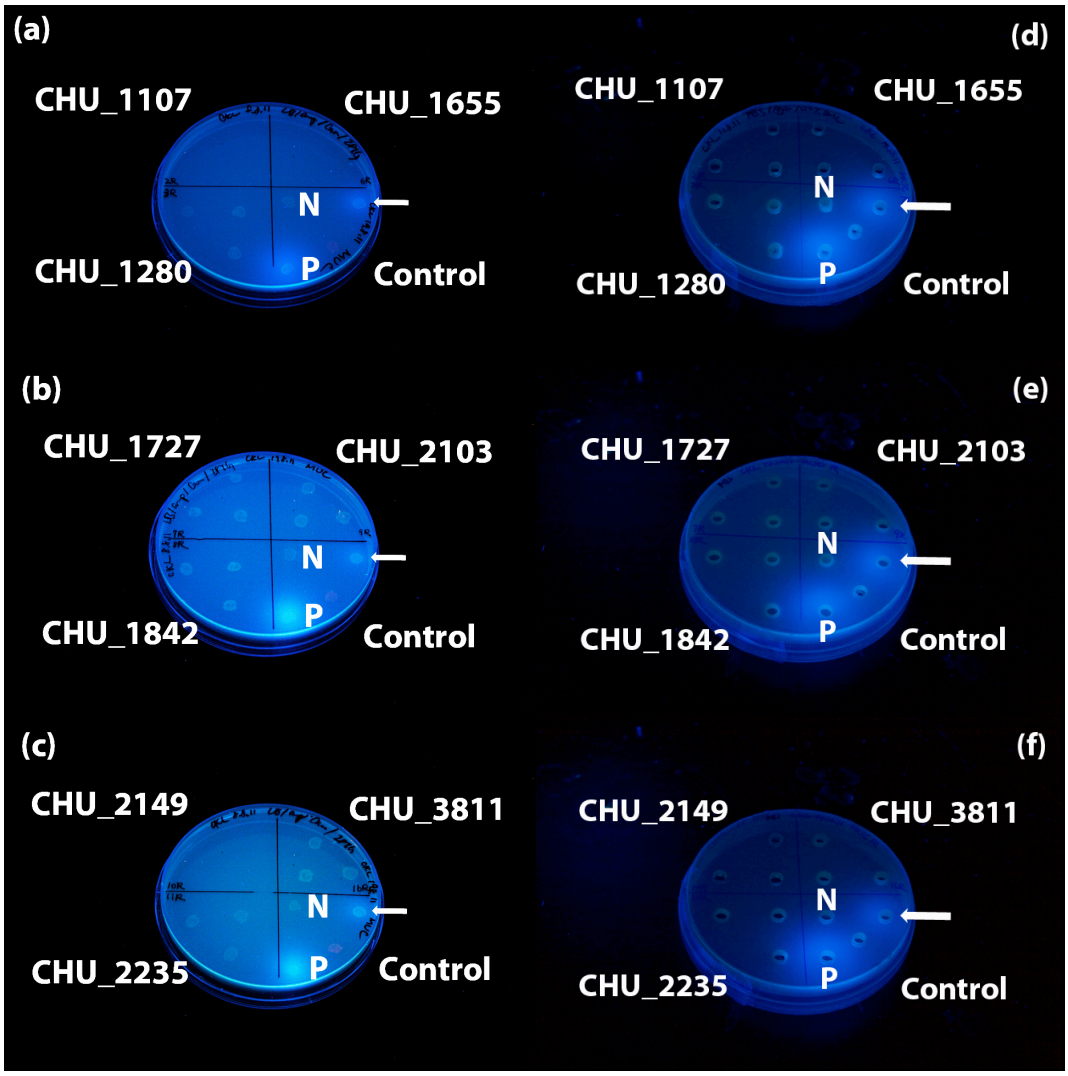


Figure 5-7. MUC assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* Rosetta. Both areas (live cells and lysate) of CHU_1107, CHU_1655 and CHU_2149 were empty due to the failure of transformation. Positive controls (P) were Cex (from *C. fimi*) and negative controls (N) were EdinBrick I vector. The un-labeled cell colonies and wells in the 'Control' areas were CenA (from *C. fimi*) and BglX (from *E. coli* MG1655). The arrows indicate CenA cell colonies or lysates that had minor MUC activities.

In Figure 5-7 (a), (b) and (c), none of the cell colonies showed MUC activity. Growth of CHU_2235 was inhibited due to the IPTG induction. CenA, indicated by the arrows in Figure 5-7 (a)(b)(c), showed the same minor fluorescence as in Section 5.2.3. The results of MUC assays with lysates are shown in Figure 5-7 (d), (e) and (f). The lysates showed no MUC activities except the positive control and CenA [indicated by arrows in Figure 5-7(d)(e)(f)]. The results in Figure 5-7 indicate that no exoglucanase activity was detected using *E. coli* Rosetta as the host strain.

5.3.4 Characterization of ß-glucosidase using *E coli* Rosetta as host strain

The ß-glucosidase activity was tested by MUG assay and was performed using both live cells and lysates (except CHU_1107, CHU_1655 and CHU_2149). The procedures and controls were the same as in Section 5.2.3.

The MUG live cells tests were performed as shown in Figure 5-8 (a), (b) and (c). All the colonies, including the positive and negative control, showed similar levels of fluorescence; however, the fluorescence was not as strong as in other fluorescence assays. Further analysis by software (Photoshop CS4) confirmed that all these colonies showed only faint fluorescence. One reason might be due to the autofluorescence of LB agar. MUG itself may also produce problems. The long incubation, usually longer than 48 h, leads to spontaneous hydrolysis in some cases and releases fluorescence (data not shown).

Lysate tests were performed as shown in Figure 5-8 (d), (e) and (f). All loaded wells clearly showed MUG activity whereas the non-loaded wells (CHU_1107, CHU_1655 and CHU_2149) did not. All negative control wells also showed MUG activities. The appearance of MUG activities in negative control

suggested that the MUG background interference in *E. coli* Rosetta seriously affected the results making it impossible to detect ß-glucosidase activity.

From the results of both MUG tests with live cells and lysates in Figure 5-8, no ß-glucosidases could be confirmed due to the fluorescent interference.

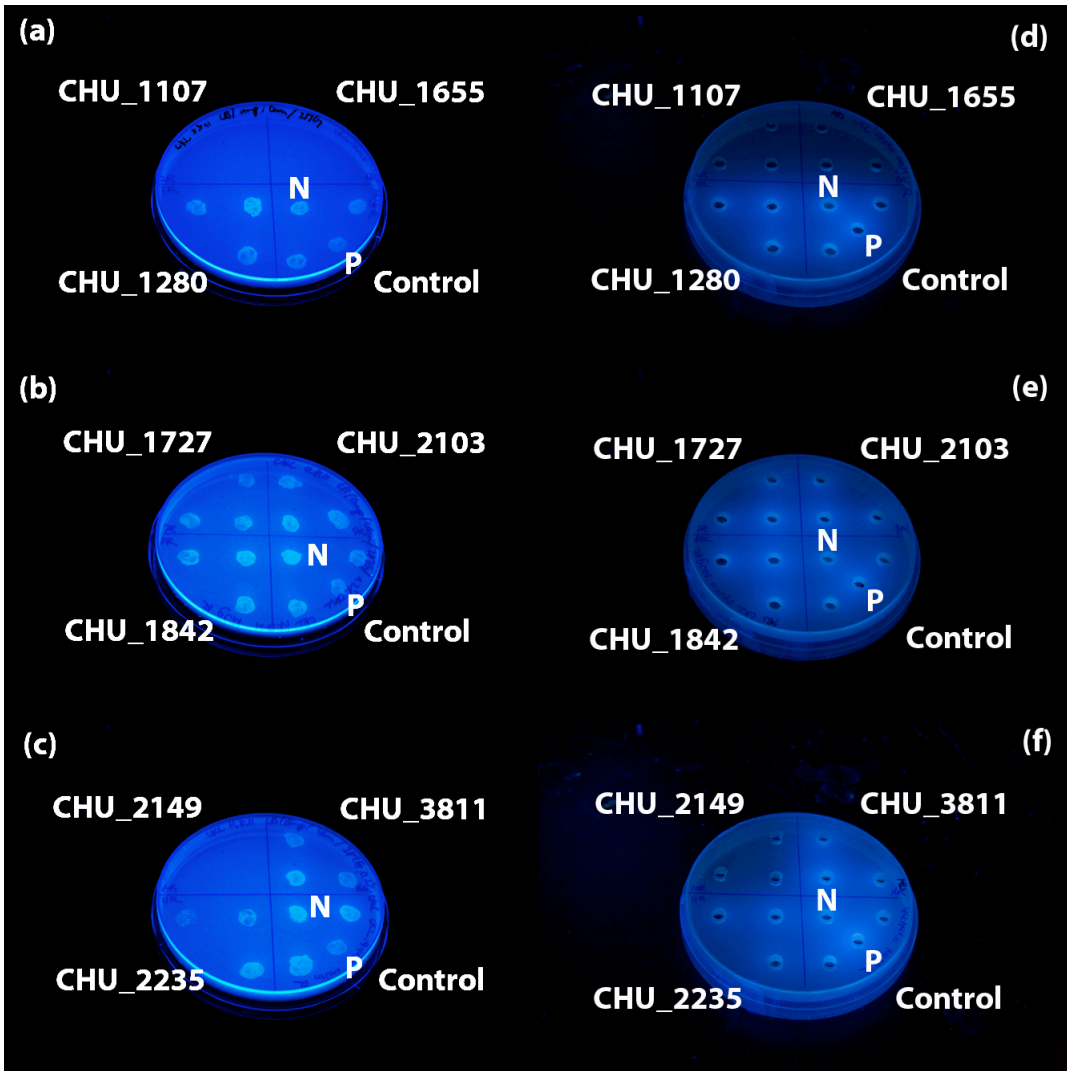


Figure 5-8. MUG assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *E. coli* Rosetta. Both areas (live cells and lysate) of CHU_1107, CHU_1655 and CHU_2149 were empty due to the failure of transformation. These figures indicate that *E. coli* Rosetta shows MUG background activity. All positive controls (P) were BglX (from *E. coli* MG1655) and negative controls (N) were EdinBrick I vector. The un-labeled cell colonies and wells in the 'Control' areas were CenA and Cex (both were from *C. fimi*).

5.4 ß-Glucosidase activity examination in *E. coli* JW2120-1 (*bglX* mutant)

As described in Section 5.2.4 and 5.3.4, MUG assays showed strong background interference. In order to reduce this interference, the *E. coli* JW2120-1 strain, which was obtained from the *E. coli* Keio Knockout Collection, was used for assays with MUG. The *bglX* gene, which is the most likely ß-glucosidase activity producer in this strain, has been disrupted by a kanamycin resistance gene.

5.4.1 Toxicity test of *E. coli* JW2120-1 expression

Expression plasmids of the nine cellulases were successfully introduced into *E. coli* JW2120-1. The procedures for toxicity tests were the same as the tests with *E. coli* Rosetta.

CHU_1280, CHU_1727, CHU_1842 and CHU_3811 showed similar cell densities, and SE variations were small. However, paired Student's t-test analysis of the mean values between induction and non-induction of CHU_1727 showed significant difference (t = -22.188, p = 0.002) even though cell densities were very similar between the induced and the non-induced cultures. The optical density of CHU_1280 was similar to CHU_1727. Its mean OD_{600} for induced treatment was a little higher than the non-induced treatment, but the statistical analysis showed no significant difference after induction (t = -3.689, p = 0.066).

CHU_1107, CHU_1655, CHU_2103 and CHU_2235 showed lower cell densities and higher SE than other cultures. Some cultures of these four genes were inhibited even without induction. One of the replicates in CHU_2149 almost failed to grow therefore contributing to the high SE. Results in Figure 5-9 suggest that some

of the constructs in *E. coli* JW2120-1 might be unstable or toxic to cells even without induction.

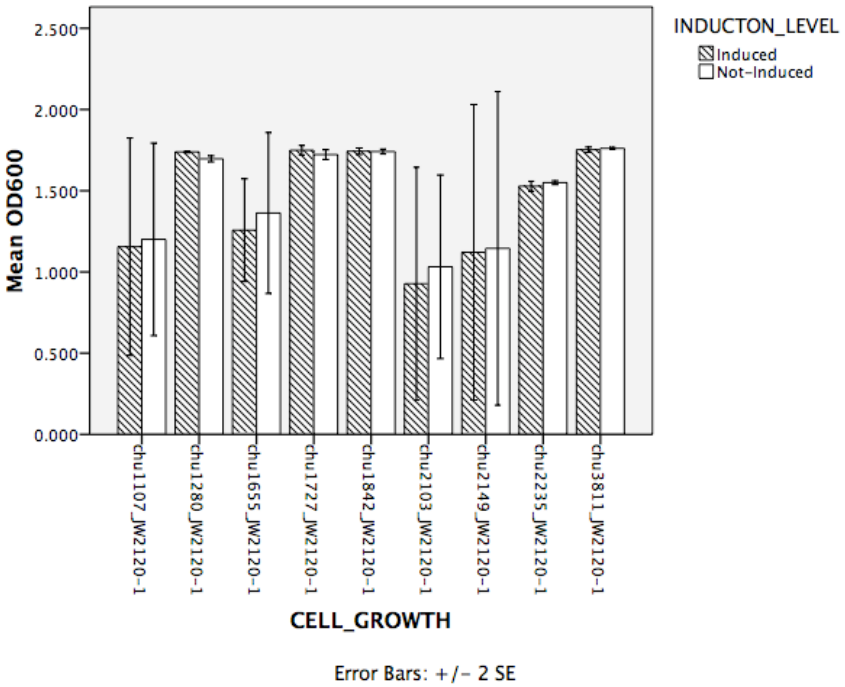


Figure 5-9. Bar chart of IPTG-induced toxicity in *E. coli* JW2120-1. CHU_1107, CHU_1655, CHU_2103 and CHU_2149 showed high SE with or without induction. The high SE suggested that plasmids of these four cellulases were possibly unstable in *E. coli* JW2120-1. Each target was cultured and measured in triplicate. The bar chart in Figure 5-9 illustrates the mean values of OD_{600} measurement after 6 h induction by IPTG (0.38 mM) and the error bar indicates ±2 SE.

5.4.2 MUG assay for ß-glucosidase using *E. coli* JW2120-1 as host strain

Results of MUG assay with live cells and crude lysates are shown as Figure 5-10. All positive and negative controls in Figure 5-10 were BglX and EdinBrick I vector, respectively. Assay procedures and preparation in Figure 5-10 were the same as in Section 5.3.4. The UV examination showed faint fluorescence in all target cells and negative control as the results in Figure 5-10 (a), (b) and (c). Only the positive control showed strong fluorescence. The growth of CHU_2235 in Figure 5-10 (c)

was too poor to verify its MUG reaction. Crude lysate tests also showed similar results. All wells of targets and negative control showed faint or no fluorescence. All positive controls showed strong blue fluorescence.

The *E. coli* JW2120-1 cultures showed significantly lower background fluorescence with MUG than the other *E. coli* strains. Some MUG live cell tests showed apparent slight MUG activity, including: CHU_1727, CHU_1842, CHU_2103, CHU_2149 and CHU_3811, but did not show the same reaction in crude lysate tests. Thus, still no ß-glucosidase activity could be confirmed among the nine target genes.

5.5 Cellulase characterization from *C. freundii* expression

According to the analysis in Chapter 4, some putative cellulase genes were predicted as extracellular or outer membrane proteins. Thus, the solubility of these proteins in the cytoplasm may be low and their presence may be harmful to host cells. Analysis in Section 5.2 and 5.3 confirmed that expression in some cases was toxic to host cells after induction. Also, *E. coli* secretes proteins poorly as it lacks the main terminal branch of the general secretary pathway (Type II secretion system). In this section, *C. freundii* was applied as an alternative host strain in the hope that it would secrete proteins more efficiently and thus reduce the toxicity produced during expression.

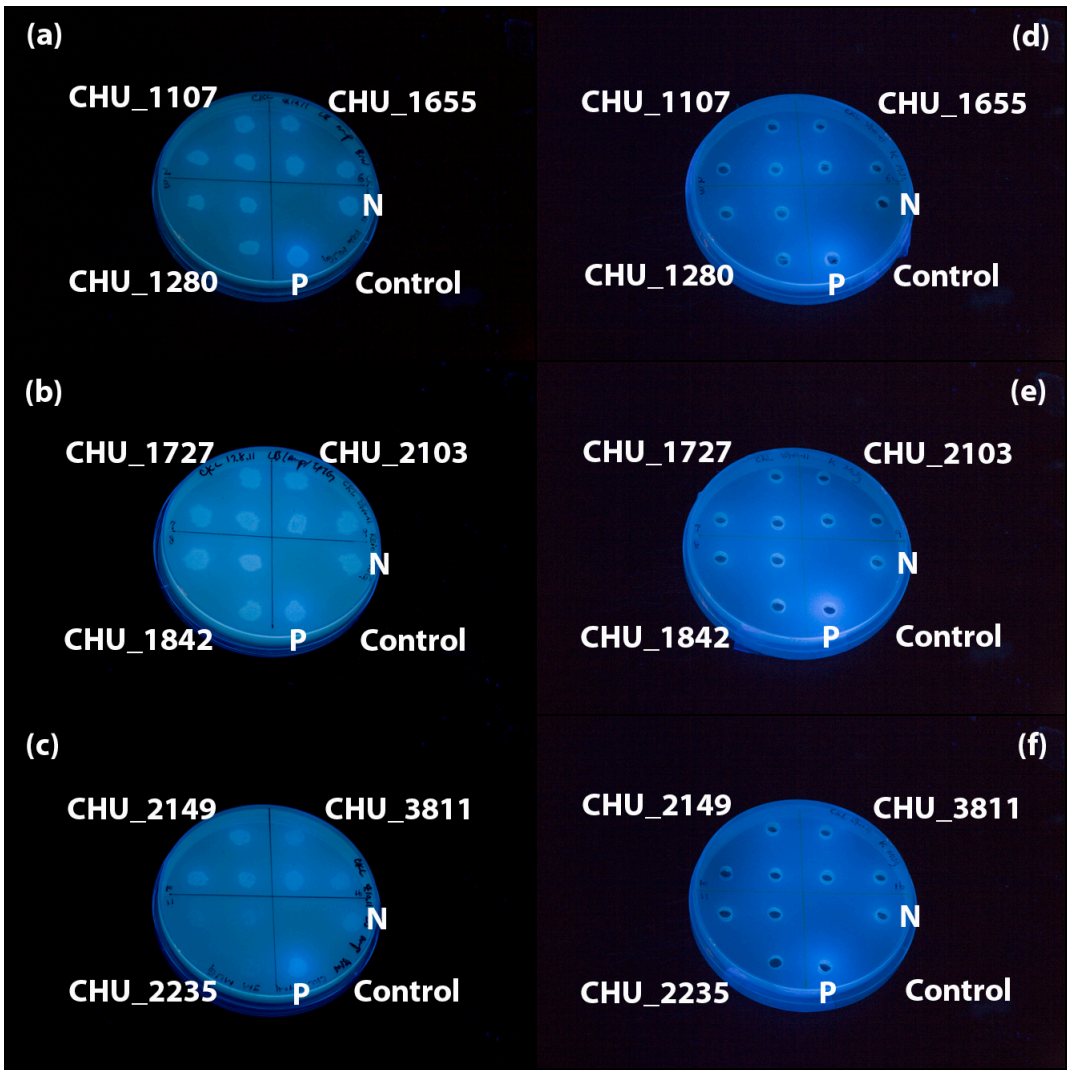


Figure 5-10. MUG assay of crude lysate for enzymes expressed in *E. coli* JW2120-1. The live cell tests are shown in (a)(b)(c), whereas lysate tests are shown in (d)(e)(f). The growth of CHU_2235 in (c) was poor possibly due to toxic expression. BglX (from *E. coli* MG1655) was applied as the positive control (P) and EdinBrick I vector was applied as negative control (N).

5.5.1 Toxicity test for *C. freundii* expression

All expression constructs were introduced into *C. freundii* competent cells successfully. Procedures of toxicity tests of *C. freundii* transformants were the same as the tests of *E. coli* Rosetta. The measurements of cell density were calculated and

graphed as shown in Figure 5-11. Paired Student's t-test of all targets showed no significant differences between the induced and non-induced treatments indicating that IPTG induction caused no negative effects in *C freundii* expression.

The optical densities in Figure 5-11 showed high variability. Part of the reason was due to floc formation as shown in Figure 5-12. Small flocs (about 1 ~ 2 mm long) were formed during some of the toxicity tests, especially during the first 4 ~ 6 h. The flocs could be found even in the 5 ml overnight starter culture. The flocs did not aggregate, and disappeared in some cultures when the cell density was high (about 1.0). Nonetheless, it was confirmed that the flocculation did cause false negative data of optical density.

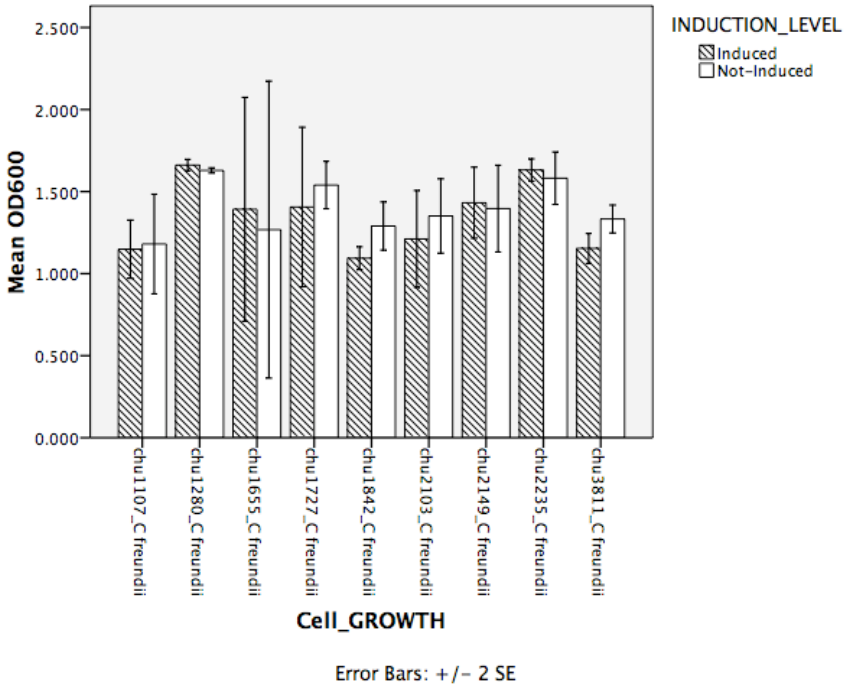


Figure 5-11. Bar chart of IPTG-induced toxicity in *C. freundii*. The mean OD_{600} measurements of the nine targets did not show growth-inhibition or suppression clearly. Part of the reason was floc formation in the cultures as shown in Figure 5-12. All tests in Figure 5-11 were prepared in triplicate and the error bars indicate ±2 SE.

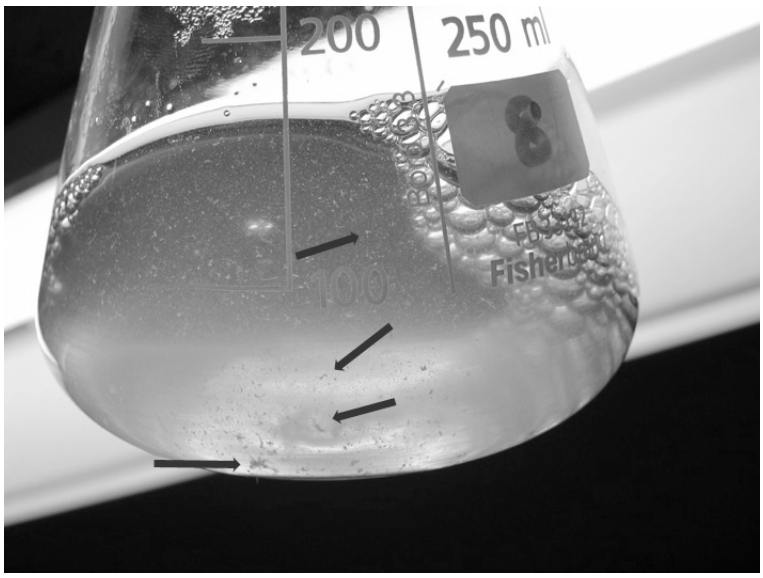


Figure 5-12. Floc formation of *C. freundii*. The floc formation (indicated by the arrows) aggregated cells and caused reduced cell optical density measurements.

5.5.2 Characterization of ß-1,4-endoglucanase activity using *C. freundii* as host strain

The preparation of CMC-Congo Red assay followed the same procedures as in *E. coli* JM109 tests (Section 5.2.2) except that the cells were cultured with carbenicillin rather than ampicillin. The results of ß-1,4-endoglucanase activity test with live cell cultures are shown in Figure 5-13 (a), (b), (c). Some of the colonies were washed away in Figure 5-13 (a), (b), (c) by sodium chloride; however, the empty area of CHU_2149 was because of the poor cell growth. Poor cell growth was also seen in the cultures of CHU_1280 and CHU_2235. This suggested that the plasmids in these cells might not be stable. One colony of CHU_1842 [Figure 5-13(b)] showed a zone of clearing indicating this enzyme might be a ß-1,4-endoglucanase. The lysate test results are shown in Figure 5-13 (d), (e), (f). The results of CHU_1842 showed much clearer and larger zones of clearing than the

negative control (EdinBrick I). Another ß-1,4-endoglucanase candidate, CHU_1280, also showed clearer zones of clearing than the negative control.

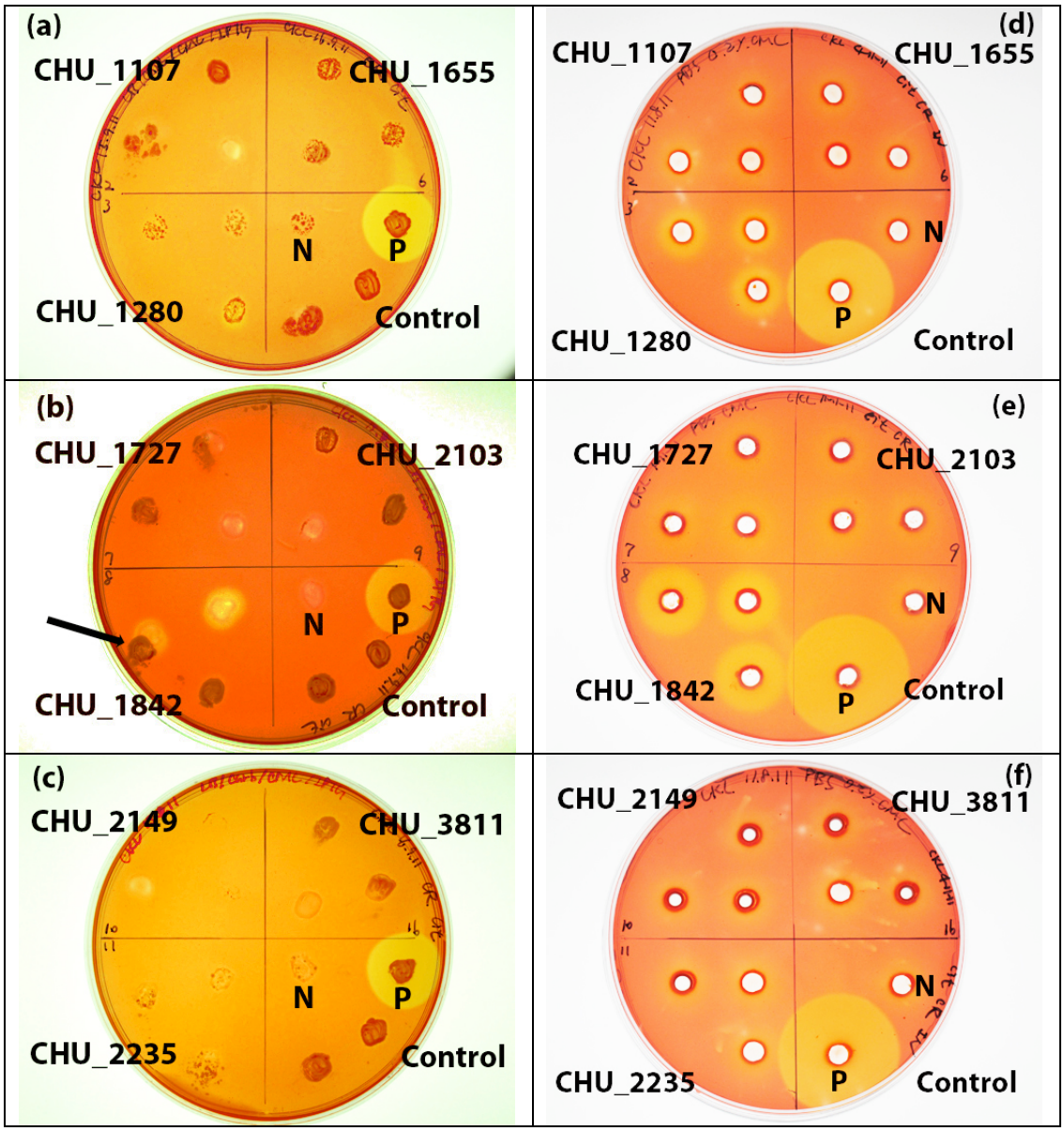


Figure 5-13. Congo Red-CMC assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] fro enzymes expressed in *C. freundii*. Positive controls (P) in both tests (cells and lysates) were CenA (from *C. fimi*) and negative controls (N) were EdinBrick I vector. In (a)(b)(c), some colonies were partially or totally washed away by 1 M sodium chloride. In (b), one CHU_1842 colony (indicated by the arrow) was washed away, so this blank area does not represent proof of cellulase secretion. In (d) and (e), lysate of CHU_1280 and CHU_1842 showed apparent ß-1,4-endoglucanase activities.

5.5.3 Characterization of exoglucanase activity using *C. freundii* as host strain

The exoglucanase activity was analyzed using MUC as substrate. Assay preparation was the same as the tests in *E. coli* JM109 (Section 5.2.3). Results of live cell and lysate tests of target genes expressed in *C. freundii* with MUC are shown in Figure 5-14.

Figure 5-14 (a), (b) and (c) represent the tests of exoglucanase activity with colonies. CHU_1280, CHU_2149 and CHU_2235 grew poorly or failed to grow. Almost all constructs, except CHU_1280 and CHU_2235, showed partial reaction with MUC but not as strong as the positive control. Comparing the exoglucanase tests with lysates in Figure 5-14 (d), (e) and (f), no target genes showed MUC activities except the positive controls. Cell debris collected from the sonication for crude lysate was further analyzed using MUC as substrate. Fluorescence was only detected in the positive controls [Figure 5-16 (a), (b) and (c)]. Hence, the partially MUC-reactive cell colonies might be false results. Ultimately, no target genes could be confirmed to possess exoglucanase activity.

5.5.4 Characterization of ß-glucosidase using *C. freundii* as host strain

ß-Glucosidase activity in *C. freundii* expression was performed using MUG as substrate. The assays were performed using the same protocols as the tests in *E. coli* JM109 (Section 5.2.4). The results of MUG assays are shown in Figure 5-15.

Figure 5-15 (a), (b) and (c) show the MUG results of colony tests, in which all positive controls failed to show fluorescence. Almost all target cells showed clear fluorescence, even those which grew poorly, such as CHU_1280, CHU_1655,

CHU_2149 and CHU_2235. However, the negative controls and the two un-labeled cell colonies (CenA and Cex) in the control area also showed strong fluorescence suggesting that *C. freundii* might produce MUG-degrading enzymes.

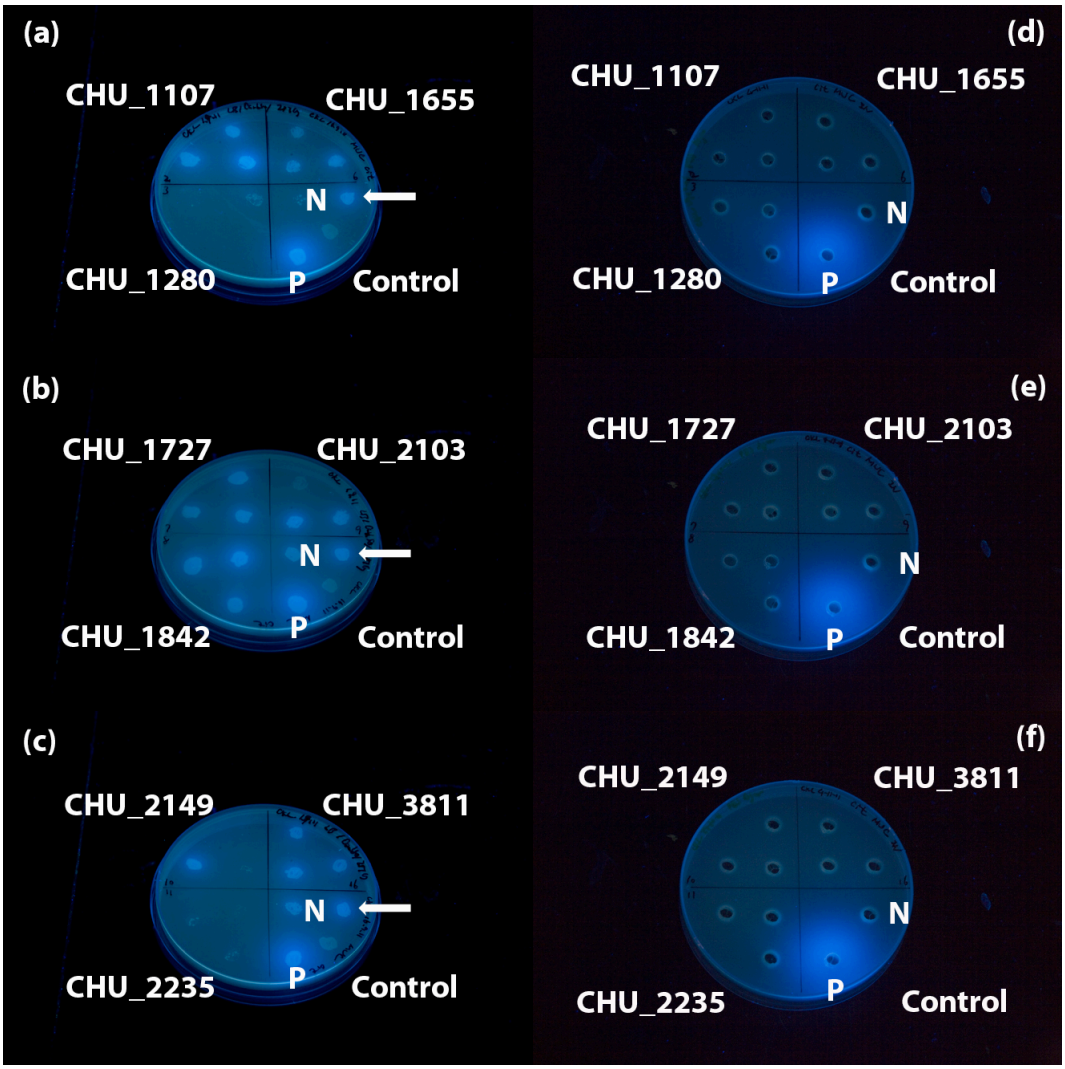


Figure 5-14. MUC assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *C. freundii*. Cex (from *C. fimi*) was applied as the positive control (P) and EdinBrick I vector was applied as the negative control (N). The un-labeled colonies in the 'Control' areas of (a)(b) and (c) are CenA (from *C. fimi*, as the arrows indicated) and BglX (from *E. coli* MG1655). Some colonies of CHU_1280, CHU2149 and CHU_2235 in (a) and (c) were missing due to poor growth.

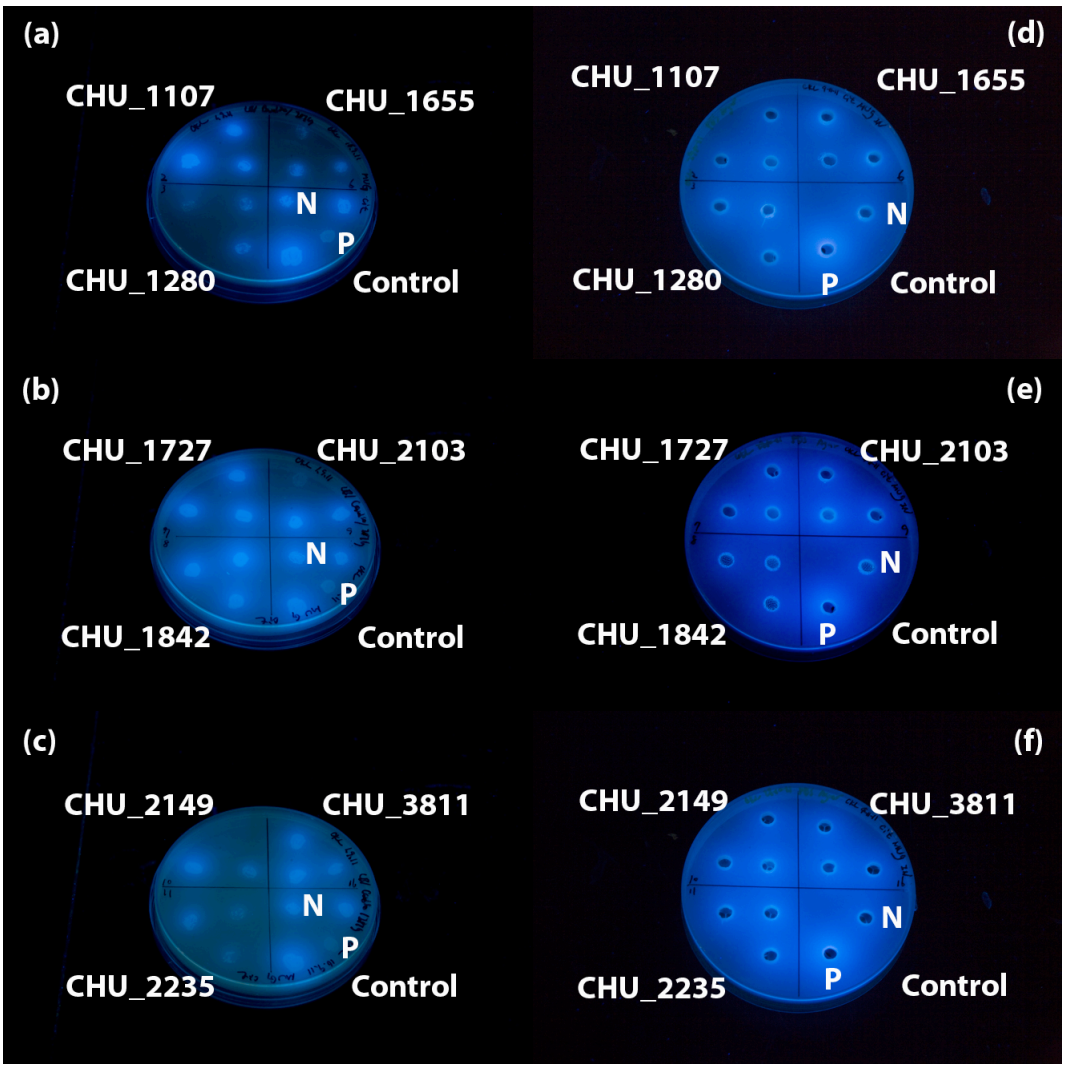


Figure 5-15. MUG assay of live cells [(a), (b), (c)] and crude lysate [(d), (e), (f)] for enzymes expressed in *C. freundii*. BglX (from *E. coli* MG1655) was used as the positive control (P) and EdinBrick I vector was used as negative control (N). In (a)(b)(c), un-labeled colonies in the 'Control' areas were CenA and Cex (both from *C. fimi*).

Figure 5-15 (d), (e) and (f) represent crude lysate tests with MUG. Almost all target lysates showed clear fluorescence including the negative control. The positive controls in Figure 5-15 (d), (e) and (f) revealed strong fluorescence. The cell debris from sonication was also tested with MUG as shown in Figure 5-16 (d), (e) and (f), in which only positive controls showed fluorescence.

Based on the results in Figure 5-15 and Figure 5-16 (d), (e) and (f), no ß-glucosidase activities were observed in cell debris. Due to the MUG background in the lysate tests, ß-glucosidase activity could not be identified by expression in *C. freundii*.

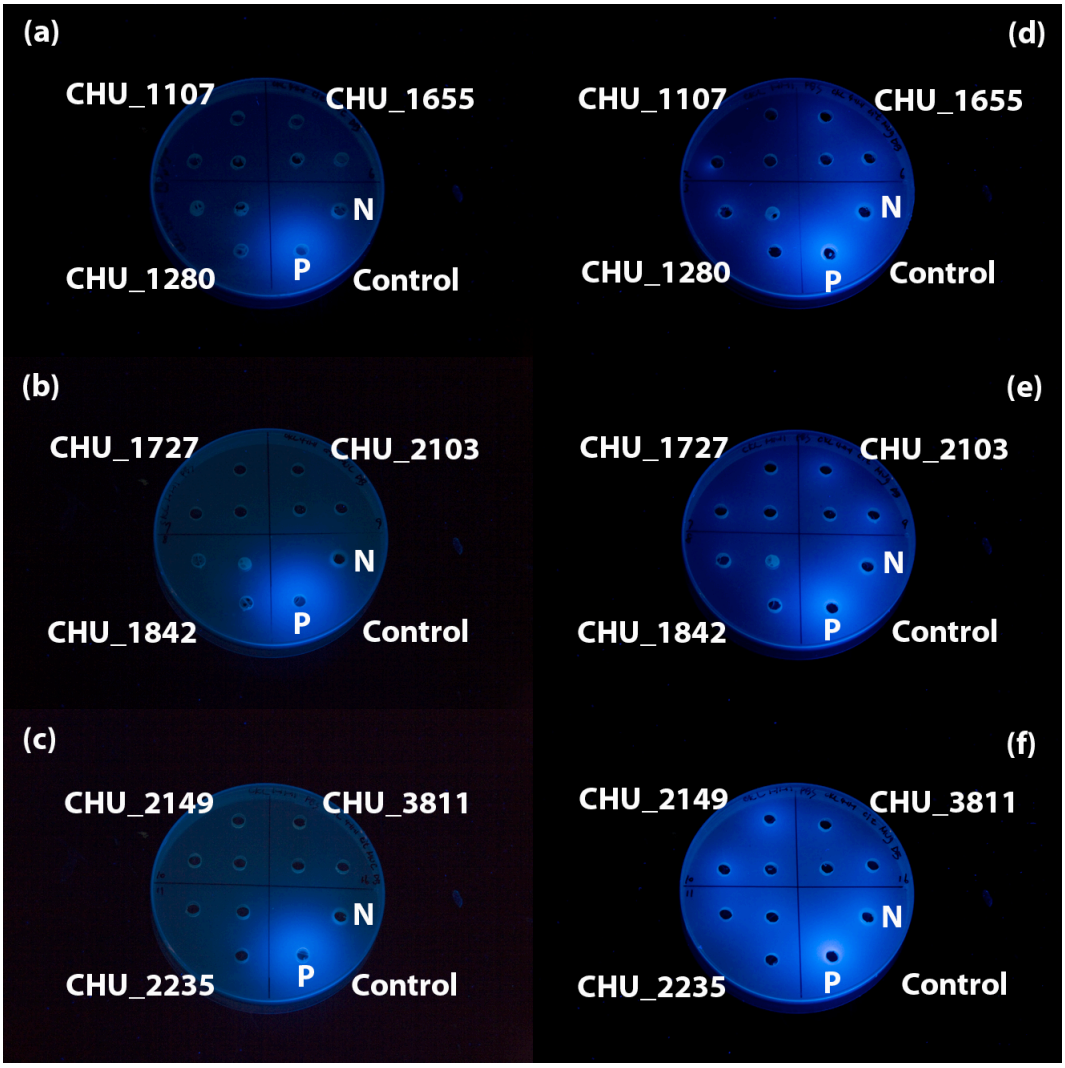


Figure 5-16. MUC [(a)(b)(c)] and MUG [(d)(e)(f)] assay of cell debris for enzymes expressed in *C. freundii*. All negative controls (N) in this Figure were EdinBrick I vector. Positive controls (P) in (a)(b)(c) and (d)(e)(f) were Cex (from *C. fimi*) and BglX (from *E. coli* MG1655), respectively.

5.6 Discussion

5.6.1 Conclusion of characterization results

In *E. coli* expression, no cellulase activities were detected with the MUG, MUC and CMC Congo Red assays. Some assays could not be done due to toxicity. However, there were some interesting results from the expression in *C. freundii*. First, CHU_1280 and CHU_1842 revealed ß-1,4-endoglucanase activities by CMC-Congo Red assay, although the activities were not as high as in the positive control. CHU_1280 and CHU_1842 are both predicted as ß-1,4-endoglucanase with unknown localizations as described in Chapter 4. BLAST results showed that CHU_1280 is related to other genes in organisms of the family *Cytophagaceae*, such as *Dyadobacter fermentans* DSM 18053 (*E* value = 0.0) and *Spirosoma linguale* DSM 74 (*E* value = 0.0). *D. fermentans* produces uncharacterized slime polysaccharide and *S. linguale* produces various different sugars, but neither species can use cellulose as carbon source [Chelius & Triplett 2000, Lail et al. 2010]. CHU_1280 is also closely related to CHU_1655 (*E* value = 0.0). All these related genes lack cellulose-binding modules. CHU_1280 also showed similarities to cellulases in *Fibrobacter succinogenes* S85 and one cellulase in *Cellvibrio japonicus* Ueda107. A gene from an un-cultured organism sample from the gut content of two earthworms, *Aporrectodea caliginosa* and *Lumbricus terrestris,* revealed high similarity (*E* value = 0.0) with CHU_1280. This cellulase was determined as a ß-glucosidase by Beloqui et al. (2010). This ß-glucosidase showed strong activity with *p*-nitrophenyl(PNP)-ß-D-glucopyranoside. However, according to their research, this

cellulase also showed higher enzyme activity on CMC than on PNP-ß-D-glucopyranoside.

CHU_1842 showed fewer BLAST results. The top three matches were from *Plesiocystis pacifica* SIR-1 (*E* value = 7×10^{-143}), *Teredinibacter turnerae* T7901 (*E* value = 9×10^{-80}) and *Arthrobacter phenanthrenivorans* Sphe3 (*E* value = 2×10^{-73}). Most of the matched genes have not been characterized. The DNS method was also used to determine the specific activity of CHU_1280 and CHU_1842, but the protein expression in *C. freundii* was too low and assay results were below the detection limit (data not shown). The CMC degradation of CHU_1280 and CHU_1842 also indicated that these two cellulases possessed hydrolysis ability even without a binding module.

There is another interesting result from the MUC tests in *C. freundii* expression as shown in Figure 5-14 and Figure 5-16. Some subcultured colonies with MUC activities did not correspond to the lysate and cell debris tests. According to the research of Kämpfer et al. (1991), MUC background should not be produced in *C. freundii* (discussed in the following paragraph). Hence, the positive MUC activities in Figure 5-14 might be true. The MUC test with cell cultures of *C. freundii* had already been performed many times in this project, but only Figure 5-14 showed positive results. Additionally, colonies for live cell subculture were picked randomly and were not the same colonies as for the preparation of lysate and cell debris. Thus, the positive results of MUC assay with live cells of *C. freundii* (Figure 5-14) may suggest that this microorganism is able to harbor and express these plasmids but that they are unstable.

The final interesting result from *C. freundii* expression is the MUG background interference. In Figure 5-15, MUG assays failed due to strong background interference. Although no ß-glucosidase has been characterized from *C. freundii*, related activity has been mentioned by Perry et al. (2007) and Kämpfer et al. (1991). Perry et al. (2007) used different substrates to test ß-glucosidase activity and *C. freundii* showed positive reactions on alizarin/iron-ß-D-glucopyranoside, alizarin/aluminum-ß-D-glucopyranoside, 5-bromo-3-indoxyl-ß-D-glucopyranoside and 6-chloro-3-indoxyl-ß-D-glucopyranoside plates. Kämpfer et al. (1991) tested fourteen 4-MU-linked substrates on 47 species in the family *Enterobacteriaceae*. Their results suggest that *C. freundii* is able to act with 4-MU-α-L-arabinopyranoside (95% positive results), 4-MU-α-D-galactopyranoside (95% positive results), 4-MU-ß-D-galactopyranoside (93% positive results), 4-MU-ß-D-glucopyranoside (84% positive results) and 4-MU-ß-D-fucopyranoside (93% positive results). The other two *Citrobacter* spp., *C. diversus* and *C. amalonaticus*, also showed positive reactions (100% and 80%, respectively) with 4-MU-ß-D-glucopyranoside.

5.6.2 Background interference of MUG assays in *E. coli* strain

As shown in the MUG assay experiments, a background fluorescence was produced in *E. coli* JM109 and *E. coli* Rosetta strain indicating that these two strains may have ß-glucosidase-related activities. Not much research has been published concerning ß-glucosidase activity in *E. coli*. However, ß-glucuronidase (EC 3.2.1.31), has been discovered in *E. coli*, and used as an assay to detect *E. coli* [Thompson et al. 1990, Trepeta & Edberg 1984]. The ß-glucosidase activity in *E.*

coli may not be regularly detected and perhaps is inducible. *E. coli* NCTC 10418 was tested by 12 different substrates for ß-glucosidase activity and only alizarin/iron plate showed positive results [Perry et al. 2007]. Another glycoside hydrolase test was performed using 43 strains of *E. coli* with eleven 4-MU-related substrates and 16% of *E. coli* strains showed ß-glucosidase activity [Kämpfer et al. 1991]. Miskin and Edberg (1978) cultured 113 *E. coli* strains with a pre-incubation (induction) protocol. They determined the numbers of positive esculin-degrading strains after 24 h, 48 h and 72 h as 30, 62 and 64, respectively. They also compared esculin degradation between cultures with or without pre-incubation. After 24 h, the numbers of esculin-degrading strains with pre-incubation were 50 and no strains grew without pre-incubation. They suggested that ß-glucosidase activity in *E. coli* needs to be induced and the longer the time of incubation, the higher the ß-glucosidase activity that can be detected.

According to NCBI and other databases, several glycoside hydrolases are present in the genome of *E. coli* MG1655 and *E. coli* BL21(DE3), but only two putative cellulases have been found: one ß-glucosidase and one ß-1,4-endoglucanase. The ß-1,4-endoglucanases in strain MG1655 and strain BL21(DE3) are designated BcsC and BcsZ, respectively. Both BcsC and BcsZ are predicted as periplasmic, and have equal DNA sequence lengths, but with one different amino acid. Since 'BCS' in BcsC and BcsZ is the abbreviation of Bacterial Cellulose Synthesis, these two proteins may not be truly cellulases. However, Park and Yun (1998) analyzed BcsC using CMC, PNP-ß-D-glucopyranoside, avicel and α-cellulose as substrates and showed that BcsC in *E. coli* K12 W3110 could be secreted and reacted with CMC. This research indicates BcsC is a ß-1,4-endoglucanase as predicted and without any

ß-glucosidase activity therefore can be rejected from the list of MUG background-producing candidates.

The other MUG background-producing candidate is a putative ß-glucosidase, BglX. BglX sequences in both *E. coli* strains [*E. coli* MG1655 and *E. coli* BL21(DE3)] are almost the same except one difference in amino acid, and both are predicted as periplasmic. BglX has been characterized by Yang et al. (1996) with degrading activity on *o*-nitrophenyl(ONP)-ß-D-glucopyranoside, but not on ONP-ß-D-fucopyranoside, ONP-ß-D-xylopyranoside or PNP-N-acetylglucosamine. They also attempted to grow wild type *E. coli* and *bglX* mutants in M9 and MacConkey medium with cellobiose, but growth was poor in both cases. This suggests that BglX may be a ß-glucosidase which is not active against cellobiose.

In this chapter, *E. coli* JM109 and *E. coli* Rosetta display clear MUG background. The result from *E. coli* JW2120-1 shows that a *bglX*-mutated strain can greatly reduce MUG background activity indicating that BglX is truly a ß-glucosidase. However, the MUG-native gel tests from Nam et al. (2010) and Wright et al. (1992) showed no fluorescent band in the *E. coli* control lysate. Their results indicate no BglX or other ß-glucosidases were expressed in *E. coli*. Therefore, the mechanism of ß-glucosidase activity induction is still unknown.

Chapter 6 Characterization of CHU_2268 and Other GH Family 3 Cellulases

As shown in the lists of putative cellulases given in Chapter 4, four GH family 3 ß-glucosidases, CHU_2268, CHU_2273, CHU_3577 and CHU_3784, have similar lengths of sequences and share the same conserved domains (GH family 3 N terminal domain, GH family 3 C terminal domain and Fibronectin type III-like domain). CHU_2268 was first expressed in *E. coli* JM109 by Ms. Natasha Cain and was unexpectedly found to be active with MUC. CHU_2268, which is predicted to be a ß-glucosidase, revealed activities only with MUC, not with MUG. Attempts were made to express CHU_2268, CHU_2273, CHU_3577 and CHU_3784 in *E. coli* BL21(DE3)pLysS and *C. freundii*. The activity was determined with MUG and MUC. Expression of these enzymes was slightly toxic to *E. coli* BL21(DE3)pLysS. CHU_2273 and CHU_3784 showed activity with MUG, not with MUC. Native-PAGE analysis also clearly showed that the activity of CHU_2268, which was active with MUC only, was different from that of CHU_2273 and CHU_3784. CHU_2268 showed 10 times higher activity (3.2×10^{-3} U/mg) than CHU_2273 (2.2×10^{-4} U/mg) and CHU_3784 (2.6×10^{-4} U/mg) with MUC. All of the analyses indicated that CHU_2268 is a cellobiohydrolase (EC 3.2.1.91) rather than a ß-glucosidase (EC 3.2.1.21) or a glucan 1,4-ß-glucosidase (EC 3.2.1.74).

6.1 Preliminary test results for CHU_2268

6.1.1 Preliminary test with *E. coli* JM109 as host strain

The BioBrick-designed plasmid constructed by Ms. Cain containing CHU_2268 under the control of a *lac* promoter, was first introduced into *E. coli* JM109. The transformants were directly streaked onto LB plates with appropriate antibiotics, 0.38 mM IPTG, X-gal and the substrate (MUC or MUG) for overnight culture at 37 ºC. The results are shown in Figure 6-1 where (a) and (c) are MUC tests and (b) and (d) are MUG tests. The null controls in Figure 6-1 (c) and (d) indicate the areas with only MUG or MUC substrate. In Figure 6-1 (a) and (c), only Cex showed fluorescence and all CHU_2268 subcultures and negative controls showed no reactions with MUC. Also, the null control showed no fluorescence indicating that MUC was not spontaneously self-degraded to release 4-MU. These plates showed that no MUC activity was detected in live cell expressing CHU_2268. The same cells were also subcultured to plates with MUG and the results were as shown as in Figure 6-1 (b) and (d). The MUG tests showed unexpected results within the controls. Figure 6-1 (b) shows that no fluorescence was discovered in any CHU_2268 expressing cells, whereas Cex, CenA, EdinBrick I vector and one subculture of CHU_2268 showed fluorescence in Figure 6-1 (d). The null control in Figure 6-1 (d) showed that no fluorescence was produced therefore indicating that the MUG reactions in the four subcultures were due to the cells.

Further tests were performed with cell lysates. The transformants of *E. coli* JM109 were cultured overnight in 50 ml LB culture with antibiotics and 0.38 mM IPTG at 37 ºC. The liquid cultures were harvested by centrifugation. Supernatant was retained to test for the secreted enzyme activity, and the pellets were

lysed by sonication to prepare cell extracts. The MUC results of the lysates were as shown in Figure 6-2. The lysate and the cell debris of CHU_2268 had strong fluorescence, as did the positive control of Cex, but not the EdinBrick I vector, which was used as the negative control. This assay also showed faint fluorescence in the CHU_2268 supernatant indicating that some of the CHU_2268 expression was secreted into the culture medium.

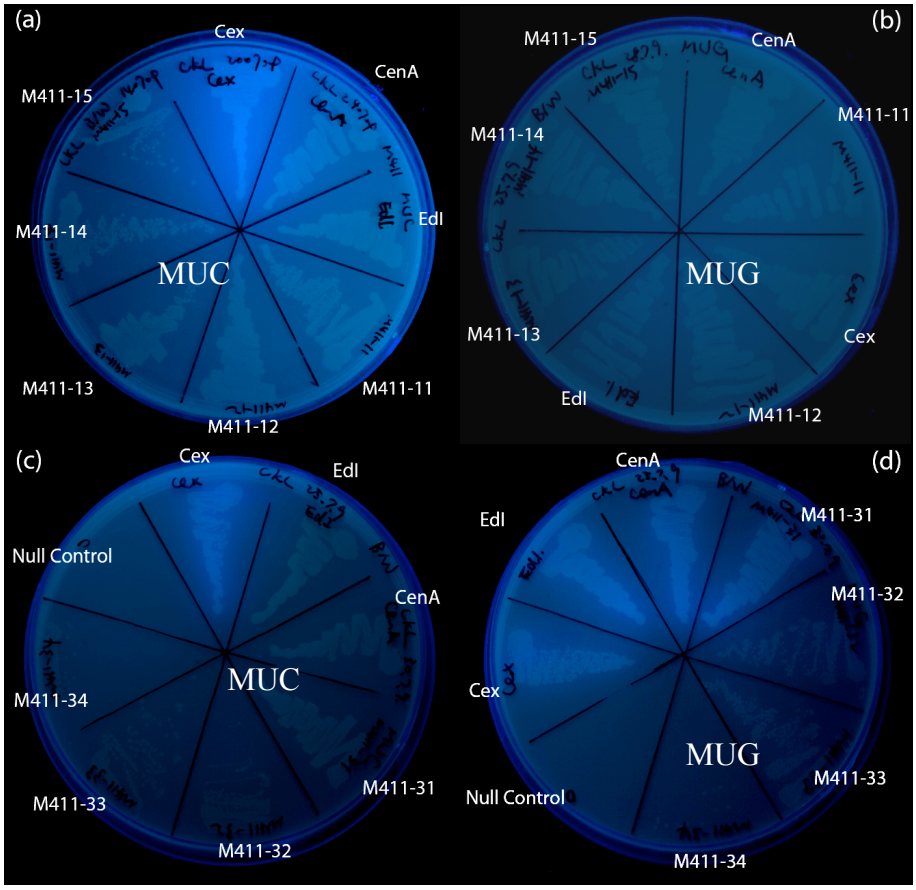


Figure 6-1. CHU_2268 live cell test with MUG and MUC as substrate. The images (a) & (c) were the tests with MUC; images (b) & (d) were the tests of MUG. These preliminary tests of CHU_2268 showed no MUC activity but with one positive result from MUG tests. However, the MUG tests showed inconsistent results within the controls. [Abbreviations: Cex: positive control for MUC; EdI: EdinBrick I vector; CenA: negative control; M411-11 to M411-34: *E. coli* JM109 transformants of CHU_2268; Null control: area with only substrate.]

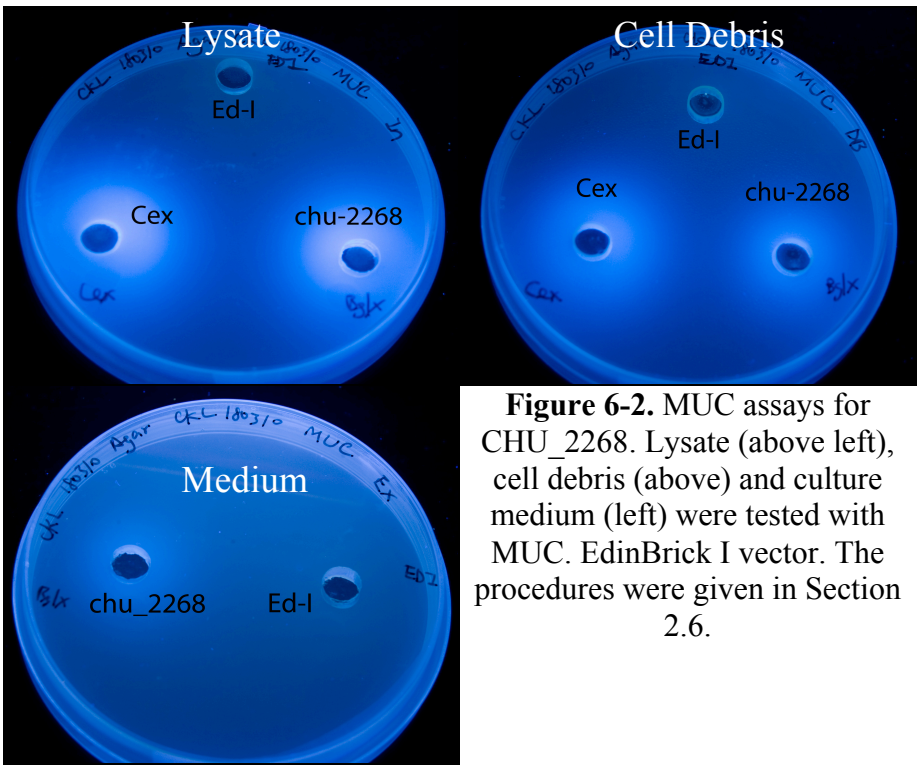


Figure 6-2. MUC assays for CHU_2268. Lysate (above left), cell debris (above) and culture medium (left) were tested with MUC. EdinBrick I vector. The procedures were given in Section 2.6.

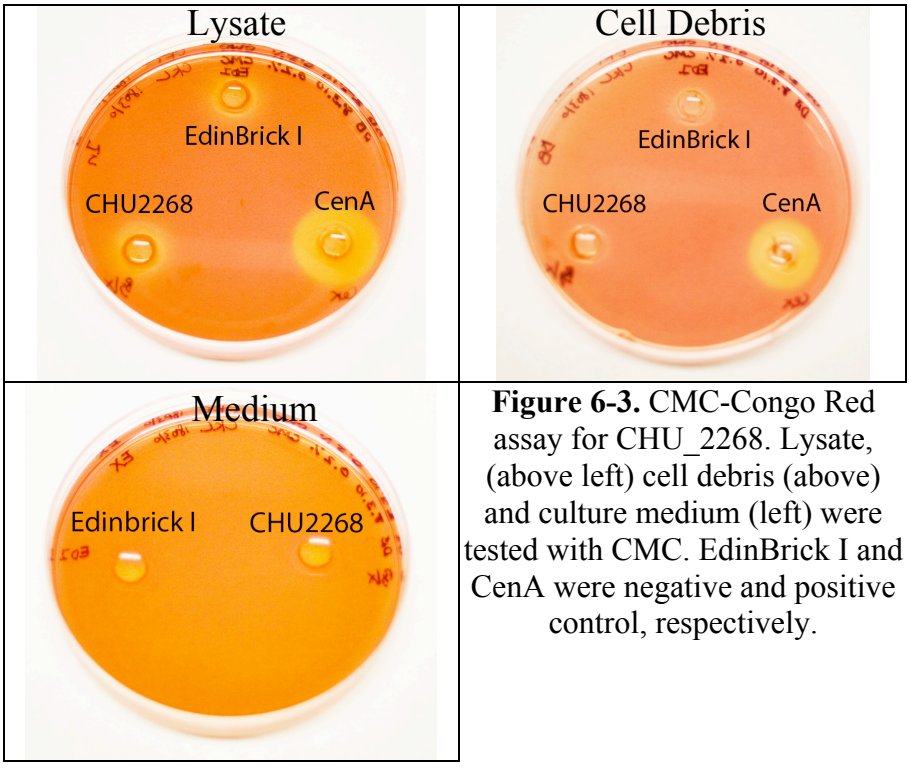


Figure 6-3. CMC-Congo Red assay for CHU_2268. Lysate, (above left) cell debris (above) and culture medium (left) were tested with CMC. EdinBrick I and CenA were negative and positive control, respectively.

The same lysates were also tested for ß-1,4-endoglucanase activity with the Congo Red-CMC assay, and the results were as shown in Figure 6-3. It showed that the CHU_2268 transformants had very low CMC degrading ability similar to the EdinBrick I vector control.

CHU_3784 was first expressed by using *E. coli* JM109, *E. coli* Rosetta and *E. coli* JW2120-1 as the host strains. However, no cellulase activity was discovered (data not shown). CHU_3784 was further expressed using *C. freundii* as the host strain and results of its enzyme activity tests were as shown in Figure 6-4. The MUG and MUC assays in Figure 6-4 were performed using the sonication products of the cell pellet harvested from a 50 ml overnight LB culture with 0.38 mM IPTG. The lysate, cell debris and supernatant were loaded onto the MUC or MUG plate for overnight incubation at 37 ºC.

The right-hand picture in Figure 6-4 shows that CHU_3784 had no MUC-hydrolyzing activity in lysate, cell debris and culture medium, whereas the left-hand picture shows that the cell debris had MUG-hydrolyzing activity. The results in Figure 6-4 suggested that CHU_3784 might have cellobiose-degrading ability and the expressed protein might be located in the periplasm or membrane-associated.

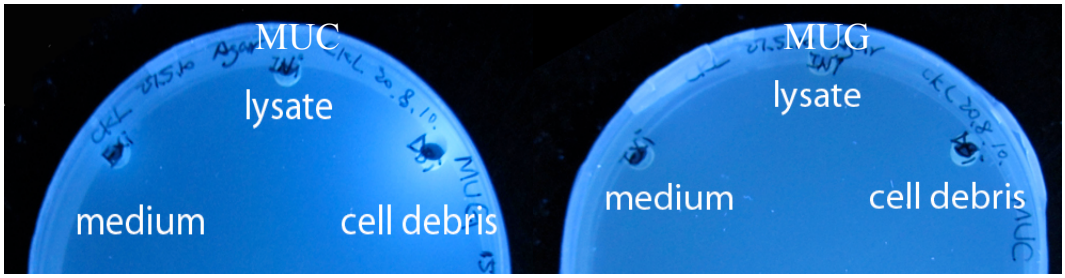


Figure 6-4. MUC and MUG tests of CHU_3784 expressed in *C. freundii*. The lysate and cell debris samples were from the sonication. Medium samples were directly obtained from the cultures of *C. freundii*. Each well was loaded with 50 µl of sample.

6.1.2 Quantitative MUG and MUC activity assay of CHU_2268

Two experiments measuring enzyme activities of the cell lysates using MUC and MUG as substrate are shown in Figure 6-5 and Figure 6-6, respectively. Four plasmids (EdinBrick I, Cex, CenA and CHU_2268) were transformed into *E. coli* JM109 and the transformants were inoculated in 50 ml LB with appropriate antibiotics and 0.38 mM IPTG for overnight incubation at 37 ºC. The crude lysates were harvested after centrifugation and sonication of the overnight culture. The absorbance of the assay was measured at 348 nm.

The results with MUC are shown in Figure 6-5. Figure 6-5 (a) represents the results with EdinBrick I vector. All four volumes of EdinBrick I lysates showed no MUC-hydrolyzing activities. The assays of CHU_2268, Cex and CenA in Figure 6-5 (b), (c) and (d) showed that the absorbance proportionally increased with time and the different test volume of lysate was also in proportion to the rate of absorbance change. The results showed that CHU_2268 was expressed in *E. coli* JM109 successfully and that MUC-hydrolyzing ability of CHU_2268 could be maintained for at least 24 h. CenA, in this experiment, also showed MUC-degrading ability.

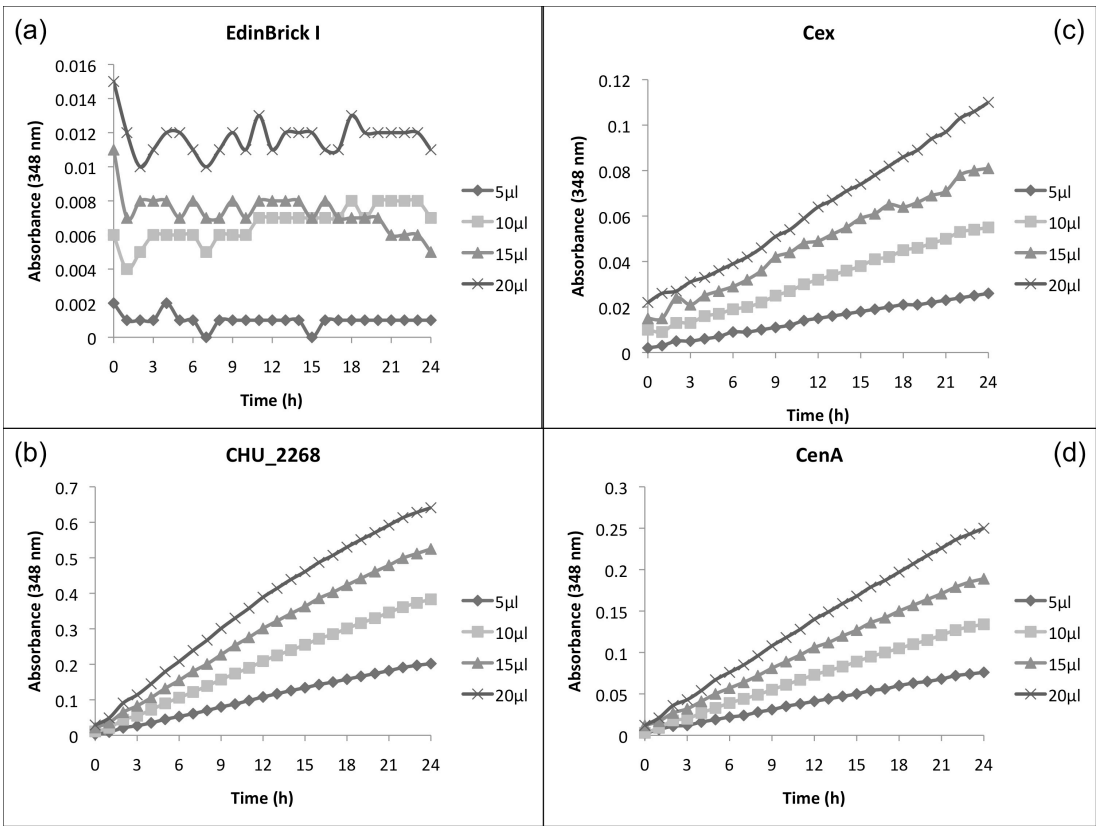


Figure 6-5. MUC activity assays. Four lysates harvested from *E. coli* JM109 (EdinBrick I, CHU_2268, Cex and CenA) were mixed with MUC. Each assay was prepared in a total 1 ml mixture containing 5, 10, 15 or 20 µl lysate, 0.2 mM MUC and PBS (pH 7.0±0.2). Assays were incubated at 37 ºC for 24 h. The release of 4-MU was measured every 1 h at 348 nm.

Similar enzyme activity assays were also performed with MUG and the results are shown as in Figure 6-6. The activity seen with Cex, CenA and CHU_2268 was only slightly higher than the negative control, EdinBrick I vector. These four lysates showed very low MUG-hydrolyzing ability compared to the MUC assay as shown in Figure 6-5.

6.2 T7 promoter expression for enzyme assay tests

The four putative GH family 3 ß-glucosidases of *C. hutchinsonii* were inserted into pT7-7 for T7 promoter expression. The primers are listed in Appendix

II and constructed plasmids were checked by sequencing (The GenePool, University of Edinburgh). The plasmids were then introduced into *E. coli* BL21(DE3)pLysS. The transformation was repeated several times and the transformation efficiencies were all very low. The numbers of transformants of all four putative ß-glucosidases were very few, usually less than 10 colonies per plate. The low transformation efficiency suggested that these four plasmids might be toxic to the host cells.

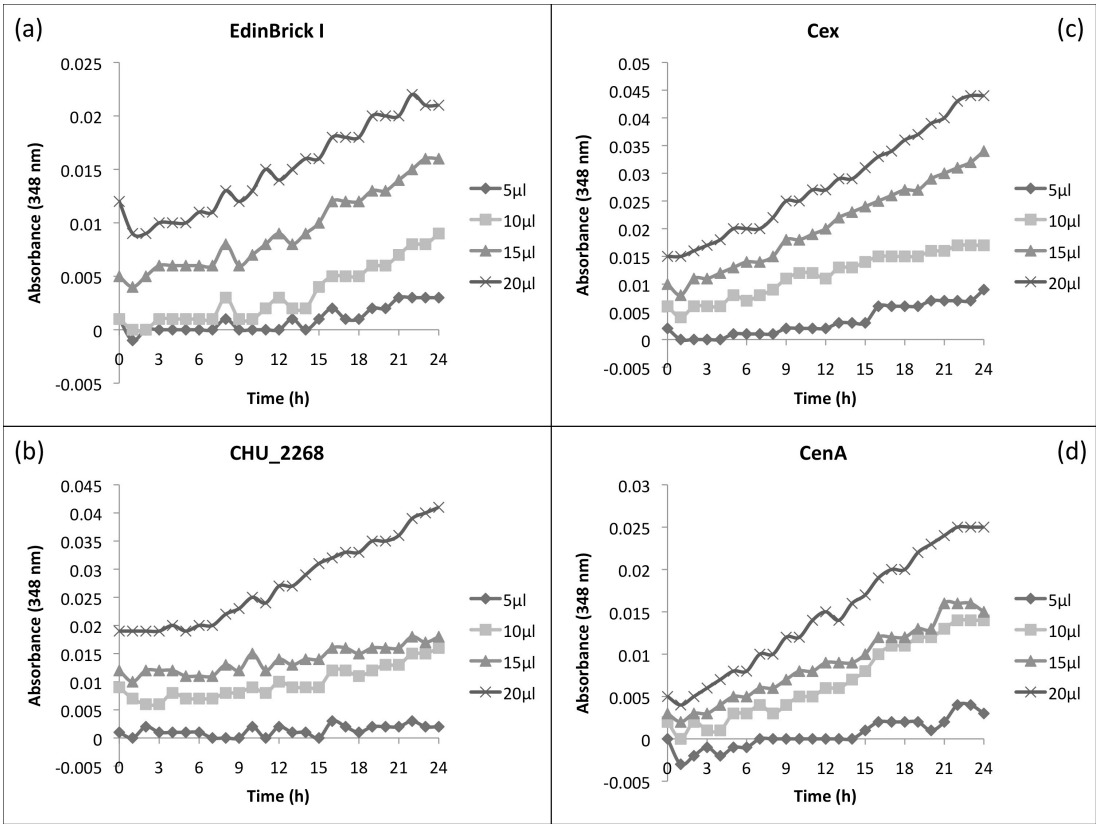


Figure 6-6. MUG activity assays. Four lysates harvested from *E. coli* JM109 (EdinBrick I, CHU_2268, Cex and CenA) were mixed with MUG and incubated at 37 ºC. Each assay was prepared in a total 1 ml mixture containing 5, 10, 15 or 20 µl lysate, 0.2 mM MUC and PBS (pH 7.0±0.2). The absorbance (348 nm) was measured every 1 h for 24 h.

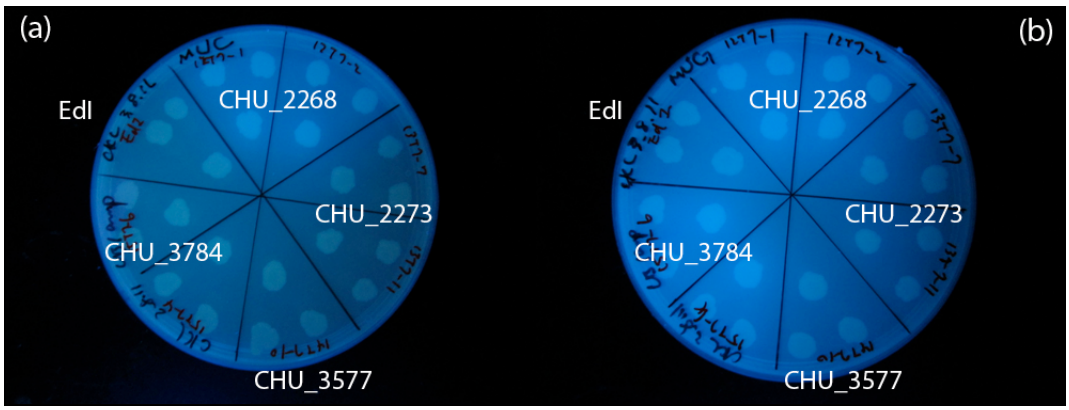


Figure 6-7. MUC and MUG assays of live cells with *E. coli* BL21(DE3)pLysS expression. (a) MUC assay and (b) MUG assay, where EdI represents the transformants of EdinBrick I vector. EdinBrick I vector and CHU_3577 were subcultured in triplicate, others used six independent transformants.

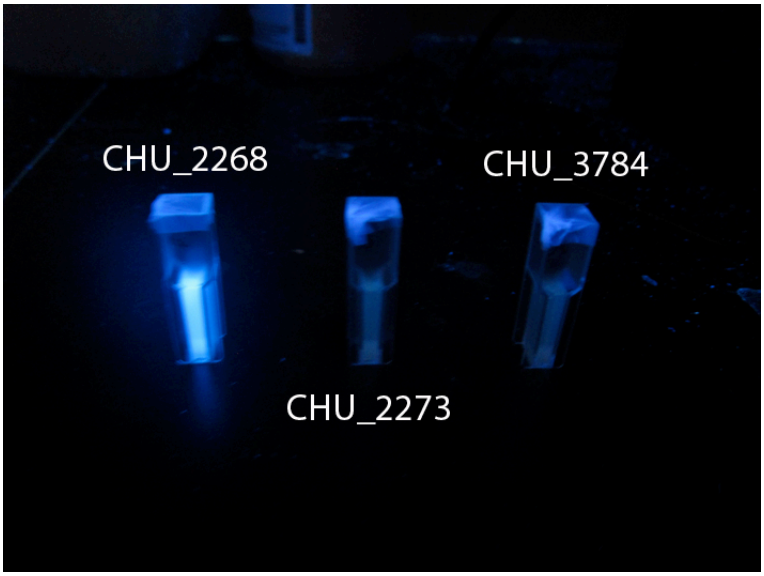


Figure 6-8. MUC assay with cell lysate from the T7 promoter expression. The image was taken under UV light (λ = 365 nm). All the cuvettes were incubated for 1 h at 37 ºC. Only CHU_2268 showed fluorescence.

6.2.1 Live cell and lysate tests

After transformation, the transformants were directly subcultured onto the MUG- and MUC-containing LB plates with IPTG. The plates were then incubated overnight at 37 ºC and the results are shown in Figure 6-7. Figure 6-7 (a) shows the MUC assay and Figure 6-7 (b) shows the MUG assay, where EdinBrick I vector was

used as negative control in both pictures. For the MUC assay, only CHU_2268 showed blue fluorescence, and the rest of the subcultures showed negative results. For the MUG assay, CHU_2268 and CHU_3784 showed positive MUG activities, whereas CHU_2273 and CHU_3577 showed negative MUG activities. The negative control in Figure 6-7 (b) showed partial positive MUG activity.

The lysates of CHU_2268, CHU_3784 and CHU_2273 were harvested from 50 ml LB culture with 0.38 mM IPTG induction for 4 h and the MUC activities were tested as shown in Figure 6-8. The assay was incubated at 37 ºC for 1 h and was performed in the cuvette with the mixed solution of 50 µl lysate, 0.2 mM MUC and 930 µl PBS (pH 7.0±0.2), giving a total volume of 1 ml. Only CHU_2268 had MUC activity, and the other two lysates showed negative results.

6.2.2 SDS- and native-PAGE test

Lysates of CHU_2268, CHU_2273 and CHU_3784 were prepared for SDS- and native-PAGE analysis. CHU_2268 was expressed in 1 l LB medium and the other two were expressed in 50 ml LB medium. The procedures of expression were as described in Section 2.4.1. Cells were then harvested by centrifugation and sonicated in PBS (pH 7.0±0.2). Protocols for SDS- and native-PAGE were as described in Section 2.8.

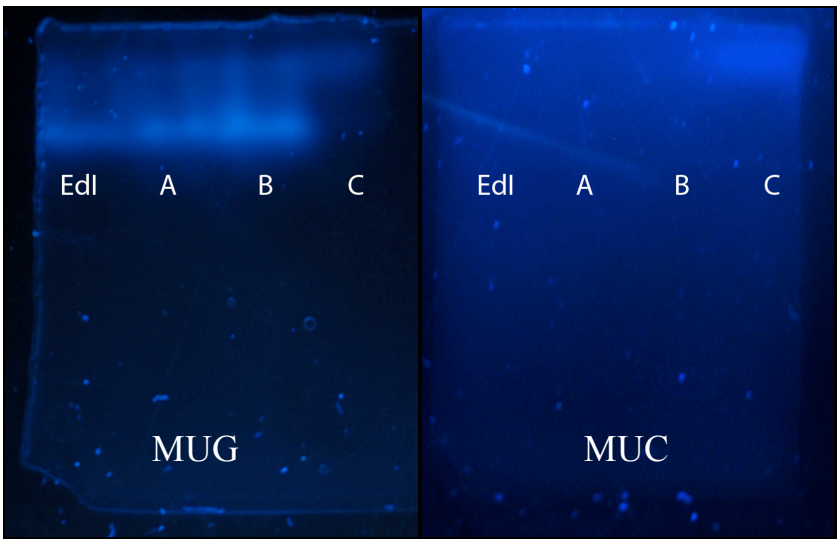


Figure 6-9. Native-PAGE test with both MUG and MUC as substrate. The left (MUG test) and right gel (MUC test) were loaded with the lysate of EdinBrick I vector (0.28 mg protein), CHU_3784 (0.31 mg), CHU_2273 (0.31 mg) and CHU_2268 (0.22 mg). Both the gels were run at the same time and all the samples were expressed in *E. coli* BL21(DE3)pLysS. Both the gels were immersed in 5 ml MUG or MUC (1 mM) for 1 h at 37 ºC. [Abbreviations: EdI: lysate of EdinBrick I vector; A: lysate of CHU_3784; B: lysate of CHU_2273; C: lysate of CHU_2268.]

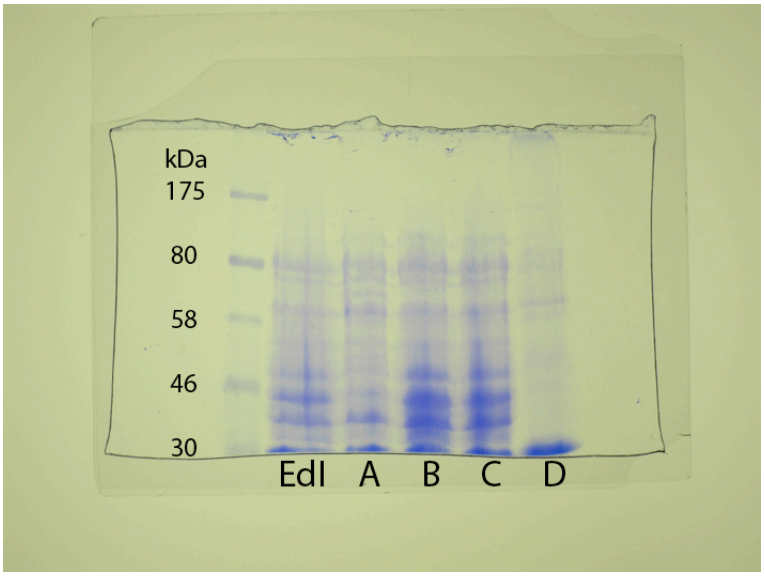


Figure 6-10. Result of SDS-PAGE with expression in *E coli* BL21(DE3)pLysS. The gel was overrun for a long time deliberately to get improved separation between 58 and 175 kDa. However, target bands still could not be verified from the SDS-PAGE. Abbreviations and the amount of sample loading: EdI: lysate of EdinBrick I vector (0.28 mg protein); A: lysate of CHU_2268 (0.22 mg); B: lysate of CHU_2273 (0.31 mg); C: lysate of CHU_3784 (0.31 mg); D: elute of CHU_2268 from ion-exchange chromatography (0.17 mg).

The native-PAGE results are shown in Figure 6-9. The lysate of EdinBrick I vector in the MUG test (left gel in Figure 6-9) showed a clear fluorescent band, and CHU_2273 and CHU_3784 also showed fluorescence at the same position indicating the background MUG activity that was produced by host cells. The sample of CHU_2268 showed a faint band in the MUG test, whereas it showed a strong fluorescence in the MUC test, and both of the bands were located at the same position in each gel. No fluorescence was seen in the MUC test, except CHU_2268.

The results in Figure 6-9 confirmed that CHU_2268 displays MUC activity unlike CHU_2273 and CHU_3784. It also showed that CHU_2268 might have slight MUG activity. The fluorescence of CHU_2273 and CHU_3784 in the MUG test gel was stronger than that seen with the EdinBrick I vector suggesting that both enzymes might also be located at that position.

Although the native-PAGE experiment showed that CHU_2268, CHU_2273 and CHU_3784 had MUG or MUC activity, no extra bands representing CHU_2268 (83 kDa), CHU_2273 (89 kDa) or CHU_3784 (81 kDa) were clearly observed by SDS-PAGE (Figure 6-10) indicating that expression levels were rather low.

6.2.3 Toxicity of CHU_2268, CHU_2273 and CHU_3784 expression

Toxicity of induction was tested in the expression of CHU_2268, CHU_2273 and CHU_3784 in *E. coli* BL21(DE3)pLysS. All the tests were performed in 50 ml LB with appropriate antibiotics. IPTG at varying concentrations was added to induce for 4 h after OD_{600} reached 1.0. The results are as shown in Figure 6-11.

All three enzymes in Figure 6-11 showed obvious growth inhibition and the decline of cell density was proportional to the volumes of IPTG. The results also indicated that the expression of these three enzymes was toxic to the host cells.

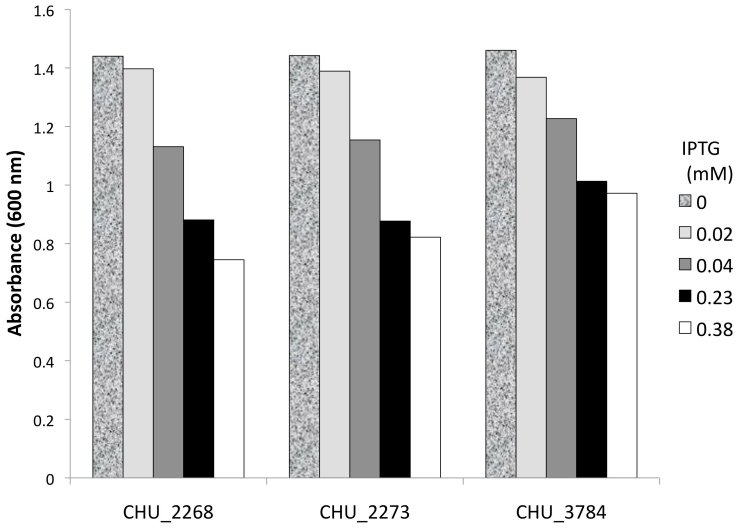


Figure 6-11. Expression toxicity test of CHU_2268, CHU_2273 and CHU_3784. The 50 ml LB culture was prepared and IPTG was added to induce for 4 h after OD_{600} reached 1.0. Each cellulase was tested with five concentrations of IPTG (0, 0.02, 0.04, 0.23 and 0.38 mM). The final OD_{600} was measured as shown in the graph.

6.3 Enzyme activity of CHU_2268, CHU_2273 and CHU_3784

The crude lysates of CHU_2268, CHU_2273 and CHU_3784 were used to determine their enzyme activities. Lysates were prepared using the same cells as described in Sec. 6.2.2 and the protocols to obtain crude lysates were as described in Sec. 2.4.1. Each enzyme activity was determined by measuring MUC or MUG hydrolysis with a mixture of 50 µl crude lysate plus 950 µl PBS (pH 7.0±0.2) and 0.28 mM MUG or 0.2 mM MUC in a cuvette with incubation at 37 ºC. The absorbance (348 nm) was measured every 10 min from 0 to 60 min and each enzyme was tested in triplicate. The measurements were then analyzed using SPSS to determine the enzyme activity with MUG or MUC.

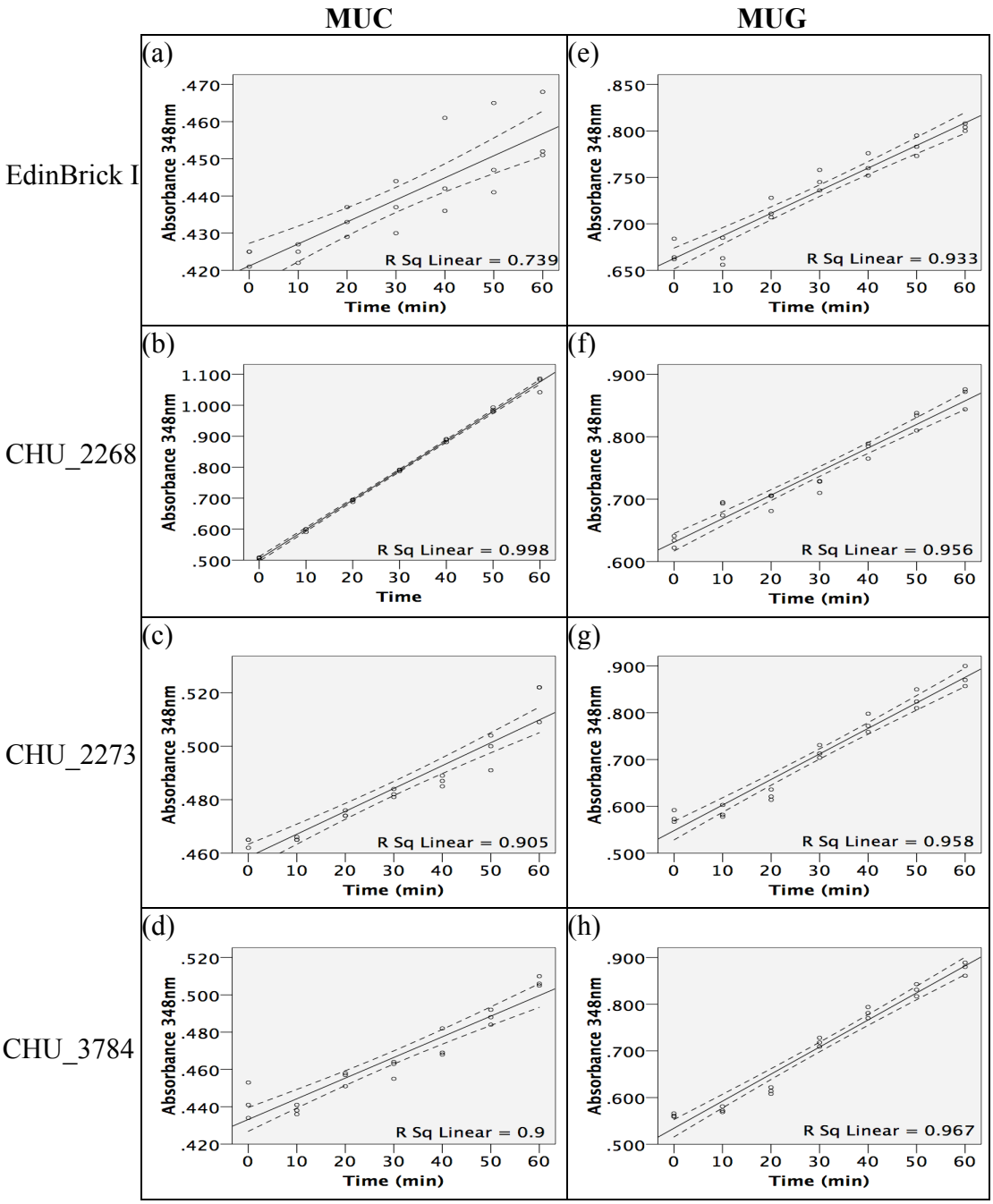


Figure 6-12. Linear regression analysis of enzyme activity tests. (a) to (d): enzyme activity tests with MUC; (e) to (h): enzyme activity tests with MUG; (a) and (e): EdinBrick I vector; (b) and (f): CHU_2268; (c) and (g): CHU_2273; (d) and (h): CHU_3784. Each assay was prepared in a mixture of 20 µl lysate, 950 µl PBS (pH 7.0±0.2) and 0.2 mM MUC or 0.28 mM MUG. The cuvettes were incubated at 37 ºC for 1 h and were measured every 10 min at 348 nm. Each lysate was tested in triplicate. The dashed line represented the 95 % confidence interval of the sample means.

Table 6-1. Linear regression analysis of enzyme activity

(a)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.860[a]	.739	.725	.007407
2	.999[a]	.998	.998	.009438
3	.951[a]	.905	.900	.005822
4	.949[a]	.900	.895	.007748
5	.966[a]	.933	.930	.013654
6	.978[a]	.956	.954	.017060
7	.979[a]	.958	.956	.024100
8	.983[a]	.967	.965	.022653

a. Predictors: (Constant), Time

(b)

Coefficients[a]

Model	Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1 (Constant)	.421	.003		144.529	.000
Time	.0006	.000	.860	7.335	.000
2 (Constant)	.504	.004		135.691	.000
Time	.095	.000	.999	92.484	.000
3 (Constant)	.459	.002		200.203	.000
Time	.0009	.000	.951	13.456	.000
4 (Constant)	.433	.003		142.121	.000
Time	.0011	.000	.949	13.096	.000
5 (Constant)	.663	.005		123.390	.000
Time	.0024	.000	.966	16.318	.000
6 (Constant)	.631	.007		94.026	.000
Time	.0038	.000	.978	20.287	.000
7 (Constant)	.548	.009		57.837	.000
Time	.0055	.000	.979	20.758	.000
8 (Constant)	.534	.009		59.953	.000
Time	.0058	.000	.983	23.462	.000

[Note: (a). R squared value prediction; (b). Regression equation prediction; Model 1 to 4: enzyme tests with MUC; Model 5 to 8: enzyme tests with MUG; Model 1 and 5: EdinBrick I vector; Model 2 and 6: CHU_2268; Model 3 and 7: CHU_2273; Model 4 and 8: CHU_3784.]

The results of linear regression analysis were as shown in Table 6-1 and the measurements are plotted in Figure 6-12. The enzyme activity of CHU_2268, CHU_2273 and CHU_3784 in MUC and MUG were calculated from the data in Table 6-1 and the results were as shown in Table 6-2. Table 6-2 also shows the specific enzyme activity against the total protein in crude lysate. CHU_2268 had the highest activity (3.2×10^{-3} U/mg) in MUC where the other lysates had about one

tenth of this activity. CHU_2268 (1.3 x 10^{-3} U/mg), CHU_2273 (1.3 x 10^{-3} U/mg) and CHU_3784 (1.4 x 10^{-3} U/mg) had similar enzyme activities in MUG tests. Lysate of EdinBrick I vector had the lowest enzyme activity in MUC (1.6 x 10^{-4} U/mg) and highest activity in MUG (6.5 x 10^{-3} U/mg) tests. All the specific enzyme activities of the lysates were low indicating that protein was expressed at a low level with *E. coli* BL21(DE3)pLysS as the host strain.

Table 6-2. Calculation of enzyme activity with different substrates

Substrate	Sample	Lysate protein concentration (mg/ml)	Δ(Abs/min)	Enzyme activity (μmol/min)	Specific activity (μmol/min/mg)
MUC	EdI	13.83	6.0×10^{-4}	2.2×10^{-3}	1.6×10^{-4}
	CHU_2268	11.02	9.5×10^{-3}	3.5×10^{-2}	3.2×10^{-3}
	CHU_2273	15.59	9.0×10^{-4}	3.4×10^{-3}	2.2×10^{-4}
	CHU_3784	15.66	1.1×10^{-3}	4.1×10^{-3}	2.6×10^{-4}
MUG	EdI	13.83	2.4×10^{-3}	8.9×10^{-3}	6.5×10^{-3}
	CHU_2268	11.02	3.8×10^{-3}	1.4×10^{-2}	1.3×10^{-3}
	CHU_2273	15.59	5.5×10^{-3}	2.1×10^{-2}	1.3×10^{-3}
	CHU_3784	15.66	5.8×10^{-3}	2.2×10^{-2}	1.4×10^{-3}

[Note: EdI: EdinBrick I vector; Abs: absorbance.]

6.4 Discussion

6.4.1 ß-Glucosidase or Exoglucanase

The four GH family 3 cellulases, CHU_2268, CHU_2273, CHU_3577 and CHU_3784, are annotated as ß-glucosidases according to their DNA sequence analysis and the annotations are all the same in the database of NCBI, KEGG, EMBL-EBI and others. However, we discovered that CHU_2268 may have MUC activity with some preliminary tests, and this activity is important since no

exoglucanases have been found in the genome of *C. hutchinsonii*. Hence, all the similar genes in *C. hutchinsonii* were searched resulting in the identification of the other three GH family 3 cellulases. As described in this chapter, these four GH family 3 cellulases have similar molecular weight (81 to 89 kDa), localization (periplasmic related) and the same putative domains. The *E. coli* BglX (see Chapter 5) is also clearly related.

Considering all enzyme tests with expression in *E. coli* strains and *C. freundii*, and the analysis of enzyme activity, CHU_2268 has consistent MUC activity and minor MUG activity. CHU_2273 and CHU_3784 show only MUG activities, whereas no activities have been found in CHU_3577. The activity results of CHU_2268 show higher MUC activity than MUG indicating that CHU_2268 releases cellobiose faster than glucose, therefore it is considered to have cellobiohydrolase (EC 3.2.1.91) activity rather than glucan 1,4-ß-glucosidase (EC 3.2.1.74) or ß-glucosidase (EC 3.2.1.21) with comparison to the mechanisms from the IUBMB database. However, no cellobiohydrolases have been characterized in GH family 3 according to the database of CAZy. No enzymes of similar sequence to CHU_2268 have been characterized by experimental measurement, except the BglX in *E. coli* MG1655 as described in Chapter 5. Nevertheless, the exoglucanase activity of CHU_2268 is still important since no exoglucanases have been found in the genome of *C. hutchinsonii*. CHU_2273 and CHU_3784 show 5 to 7 times higher activity with MUG than with MUC indicating that both of the enzymes may prefer to bind MUG rather than MUC and release glucose faster than cellobiose, therefore are both considered as ß-glucosidases. However, these results must be treated with

caution since assays were performed with crude extracts rather than purified enzymes.

Since the protein concentrations are similar in the lysates of CHU_2273, CHU_3784 (Table 6-2), the stronger MUG activities on the native-PAGE of CHU_2273 and CHU_3784 indicate both the enzymes are expressed in the lysate; also, the locations of activity bands in lysates of CHU_2273 and CHU_3784 are almost the same as that seen for the negative control (EdinBrick I vector). The similar locations of fluorescent bands in these three lysates on the native-PAGE may be because the BglX in *E. coli*, CHU_2273 and CHU_3784 are of similar size and charge. Moreover, the results from the native-PAGE analysis show background MUG activity in the negative control. However, other reported tests (native-PAGE) show no extra bands while reacting with MUG in crude lysate of *E. coli* [Nam et al. 2010, Wright et al. 1992]. More discussions of MUG activities in *E. coli* has been given in Chapter 5.

6.4.2 Binding ability of GH family 3 cellulase

All four GH family 3 cellulases discussed in this chapter have the same putative domains as described in the introduction of this chapter. These cellulases contain complete GH family 3 N- and C-terminal domains indicating that their structures should be intact and typical. Varghese et al. (1999) and Stubbs et al. (2007) studied the structure of GH family 3 hydrolases, barley ß-D-glucan glucohydrolase and NagZ enzyme, finding that their structures are globular. Varghese et al. (1999) also mentioned that the catalytic domain of barley ß-D-glucan glucohydrolase consists of a relatively shallow substrate-binding pocket that is

located at the interface of the two domains of the enzyme. Furthermore, they suggested that a second binding site might exist at the second domain (C-terminal) specific for (1-3,1-4)-ß-D-glucans. Therefore a separate substrate-binding domain in this kind of cellulase may not be necessary.

The fibronectin type III-like domains are found at the C-terminal region of these four GH family 3 cellulases. Their function in these four cellulases was not studied in this project. However, some studies suggest this type of domain may help in hydrolyzing substrates. Watanabe et al. (1994) modified a chitinase from *Bacillus circulans* WL-12 which contained a fibronectin type III-like domain and expressed it in *E. coli*. Their study showed that deletion of this domain did not affect chitin-binding activity but did result in significantly decreased colloidal chitin-hydrolyzing activity. Kataeva et al. (2002) also suggested that the fibronectin type III-like domain in the CbhA of *Clostridium thermocellum* enhanced cellulose hydrolysis as composed to the GH family 9 catalytic domain alone.

Chapter 7 Conclusions

C. hutchinsonii has been studied for almost ninety years, but its cellulose-degrading system still remains unclear. Many researchers believe this bacterium may have a unique cellulose-degrading mechanism which is different from either the free cellulase system or the cellulosome complex [Wilson 2008, Xie et al. 2007]. Since its genome was annotated in 2007, gene cloning and cellulase characterization in *C. hutchinsonii* is now much easier. Also, together with the concept of synthetic biology and BioBrick standard biological parts, simultaneous expression of multiple enzymes can be done in unified procedures.

In this project, eleven putative cellulases (including two cellulases described in Chapter 6) have been inserted into pSB1A2 using BioBrick assembly. However, in most cases, expression was toxic in *E. coli* JM109 and in the alternative strains used. The toxicity may be caused by several factors. Large numbers of rare codons in target genes may lead to problems, such as reducing translation efficiency of heterologous proteins, causing low level or undetectable expression, or mis-incorporating amino acids into the target [Chen & Texada 2006, Kane 1995]. Furthermore, *E. coli* produces numerous proteases which efficiently lyse abnormal proteins [Makrides 1996]. Proteases, such as Lon and Clp in *E. coli*, are responsible for 70-80 % of the energy-dependent degradation of proteins *in vivo* [Maurizi 1992]. Proteolysis mostly occurs during cell harvest or recovery procedures, but is normally initiated during cultivation [Bota & Davies 2002, Murby et al. 1996]. Some *E. coli* strains, such as *E. coli* BL21(DE3)pLysS and *E. coli* RF6333, are able to reduce the

negative effects of protease activity [Sambrook & Russell 2001]. In some cases, high-level expression in *E. coli*, especially for those larger or membrane-associated proteins, leads to the formation of highly aggregated complexes which are commonly referred to as inclusion bodies [Palmer & Wingfield 1995, Singh & Panda 2005]. Although enzymes can be recovered from inclusion bodies, their activity may be lost due to misfolding [Hannig & Makrides 1998]. In addition to rare codon usage and inclusion bodies, the N-terminal degron, which is a degradation signal of proteins, is also related to the *in vivo* protein stability [Varshavsky 1997]. Protein sequences with Arg, Lys, Phe, Leu, Trp, or Tyr at the N-terminus have less than 2 min half-life in *E. coli* [Tobias et al. 1991].

The expression efficiency in *E. coli* can be improved in several ways, such as optimizing rare codon usage [Gustafsson et al. 2004], modifying signal peptides, increasing protein solubility and co-expressing chaperones [Sørensen & Mortensen 2005, Weickert et al. 1996]. Although *E. coli* may not recognize the signal sequences of *C. hutchinsonii*, these sequences are best excluded from the expression plasmids. Expression of insoluble proteins may lead to the formation of inclusion bodies in *E. coli*. In this project, inclusion bodies were not seen under microscopic observation (data not shown). One of the solutions to increase protein solubility or secretion is to express a fusion protein by ligating the target sequence to a soluble or secreted protein. Chaperone co-expression may also increase the expression efficiency.

One remarkable conclusion is from the analysis of CHU_2268. *C. hutchinsonii* is well known for its capability of cellulose degradation; however, no exoglucanases have been found in its genome, leaving a mystery to be solved. Therefore, the missing exoglucanase may possibly be concealed among the

unidentified hypothetical proteins. In our experiments, CHU_2268, which is a putative ß-glucosidase, showed exoglucanase activity as described in Chapter 6. Enzymes that are similar to CHU_2268 were also expressed and tested, but no other exoglucanase activities were discovered. Although the results indicate that CHU_2268 could be a cellobiohydrolase rather than a ß-glucosidase, more direct tests, such as using cellodextrins as substrates, should be performed in the future since glucose release rate is crucial for cellulase determination [Kashiwagi et al. 1991].

As described in Chaper 1, some cellulases are able to hydrolyze multiple types of substrate. In this project, the two CMC-hydrolyzing cellulases, CHU_1280 and CHU_1842, showed no exoglucanase or ß-glucosidase activity. CHU_2268 showed only cellobiohydrolase activity; also, CHU-2273 and CHU_3784 showed ß-glucosidase activity. In addition, none of these characterized cellulases contain putative CBMs. Tomme et al. (1996) reported that cellobiohydrolases without CBM might display degrading activity on microcrystalline cellulose up to 50 % lower compared to cellulases with CBM. Unfortunately, no characterized cellulases were tested for their binding abilities in this project; hence no conclusions can be made.

Most putative ß-1,4-endoglucanases of *C. hutchinsonii* are presumably extracellular or outer membrane associated; also, CHU_2268 is predicted as periplasmic, therefore, outer membrane pores or intake apparatus should be present. Wilson (2008, 2009) emphasizes that *C. hutchinsonii* may possess a third type of cellulase system different from both the typical complexed and non-complexed forms. Xie et al. (2007) also indicate that the Sus-like proteins in *C. hutchinsonii* may play an important role in cellulose degradation. However, only a few genes in

C. hutchinsonii are related to the large substrate uptake system (Chapter 4). The predicted localization of ß-1,4-endoglucanases also suggests that these cellulases may have bi-functional cellulase activity, e.g. ß-1,4-endoglucanase combined with exoglucanase. Almost all putative CBMs detected in the genome of *C. hutchinsonii* are encoded within the xylanases, except those of CHU_1051 and CHU_1727. Also, most of these CBMs were predicted as extracellular or associated with the S-layer. Since *C. hutchinsonii* is unable to use xylose as the sole carbon source [Larkin 1989], these CBMs may improve cellulose degradation by binding hemicellulose while degrading lignocellulose. However, even with all the theoretical expectations, the exact cellulose degrading mechanism is still unclear.

The slime produced by *C. hutchinsonii* was believed to be EPS by many researchers. Martin et al. (1968) analyzed its composition from a culture with glucose as main carbon source. The original aim of EPS analysis in this project was to compare its composition in cultures grown with different carbon sources. However, this goal was not achieved. Nevertheless, TLC tests and colorimetric assays showed some interesting results (Chapter 3). Pentose was detected from the cellulose-grown cultures. Also, glucose was released during grown on cellobiose indicating that cellobiose may not be assimilated directly. Another interesting experiment done by Dr. S. Lakhundi and Dr. C. E. French (unpublished) suggested that EPS from *C. hutchinsonii* could be utilized by *Bacillus subtilis* and maintained growth of *B. subtilis* in a mixed culture with cellulose as sole carbon source.

Most of the plasmids in this project were assembled using BioBrick rules. The constructs were successful, although toxicity was observed during expression. Large-scale DNA assembly based on BioBrick design or multiple protein expression,

e.g. cellulose degradation by synergistic activity of multiple cellulases, can be developed in the future. Cellulases in *C. fimi* have been characterized by Mr. D. Barnard and Mr. S. Kane (unpublished); in addition, preliminary tests for the synergistic cellulose degradation by *C. fimi* cellulases are in progress. To extend this project from the results in hand, more EPS analysis can be performed by HPLC or other analytical procedures since cultures grown with different carbon sources may produce different EPS, based on the results of the preliminary tests. Considering the problems that may be produced during the transcriptional or translational stages when expressing recombinant proteins in *E. coli*, such as N-end rules, more BioBrick based constructs could be tried. The method of BioBrick assembly for analysis of cellulose-degrading mechanisms could also be applied to other cellulose degrading microorganisms.

References

AGARWAL, S., HUNNICUTT, D. W. & MCBRIDE, M. J. 1997. An electron microscope study of the cell surface of *Cytophaga johnsonii* and some observations on related organisms. *Proceedings of the National Academy of Sciences USA,* 94**,** 12139-12144.

AKIN, D. E. 2007. Grass lignocellulose: strategies to overcome recalcitrance. *Applied Biochemistry and Biotechnology,* 137-140**,** 3-15.

ANDERSEN, J. M., BARRANGOU, R., ABOU HACHEM, M., LAHTINEN, S., GOH, Y. J., SVENSSON, B. & KLAENHAMMER, T. R. 2011. Transcriptional and functional analysis of galactooligosaccharide uptake by lacS in *Lactobacillus acidophilus*. *Proceedings of the National Academy of Sciences USA,* 108**,** 17785-17790.

ANDRIANANTOANDRO, E., BASU, S., KARIG, D. K. & WEISS, R. 2006. Synthetic biology: new engineering rules for an emerging discipline. *Molecular System Biology,* 2**,** 2006. 0028.

ATLAS, R. M. 2010. *Handbook of microbiological media,* U.S, CRC Press Inc.

AUGUSTINE, N. R. 1976. Technology transfer from military requirements to public need. *In:* ELMER L. GADEN, J., MANDELS, M. H., REESE, E. T. & SPANO, L. A. (eds.) *Biotechnology and Bioengineering Symposium: Enzymatic Conversion of Cellulosic Materials: Technology and Application.* New York: John Wiley & Sons, Inc.

BAKER, M. G., LALONDE, S. V., KONHAUSER, K. O. & FOGHT, J. M. 2010. Role of extracellular polymeric substances in the surface chemical reactivity of *Hymenobacter aerophilus*, a psychrotolerant bacterium. *Applied and Environmental Microbiology,* 76**,** 102-109.

BARR, B. K., HSIEH, Y.-L., GANEM, B. & WILSON, D. B. 1996. Identification of two functionally different classes of exocellulases. *Biochemistry,* 35**,** 586-592.

BAYER, E. A., BELAICH, J.-P., SHOHAM, Y. & LAMED, R. 2004. The Cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. *Annual Review of Microbiology,* 58**,** 521-554.

BAYER, E. A., CHANZY, H., LAMED, R. & SHOHAM, Y. 1998a. Cellulose, cellulases and cellulosomes. *Current Opinion in Structural Biology,* 8**,** 548-557.

BAYER, E. A., SHIMON, L. J. W., SHOHAM, Y. & LAMED, R. 1998b. Cellulosomes—structure and ultrastructure. *Journal of Structural Biology,* 124**,** 221-234.

BELOQUI, A., NECHITAYLO, T. Y., LO´PEZ-CORTE´S, N., GHAZI, A., GUAZZARONI, M.-E., POLAINA, J., STRITTMATTER, A. W., REVA, O., WALICZEK, A., YAKIMOV, M. M., GOLYSHINA, O. V., FERRER, M. & GOLYSHIN, P. N. 2010. Diversity of glycosyl hydrolases from cellulose-depleting communities enriched from casts of two earthworm species. *Applied and Environmental Microbiology,* 76**,** 5934-5946.

BHASKAR, P. V. & BHOSLE, N. B. 2006. Bacterial extracellular polymeric

substance (EPS): a carrier of heavy metals in the marine food-chain. *Environmental International,* 32**,** 191-198.
BHATIA, Y., MISHRA, S. & BISARIA, V. S. 2002. Microbial β-glucosidases: cloning, properties, and applications. *Critical Reviews in Biotechnology,* 22**,** 375-407.
BOKINSKY, G., PERALTA-YAHYA, P. P., GEORGE, A., HOLMES, B. M., STEEN, E. J., DIETRICH, J., LEE, T. S., TULLMAN-ERCEK, D., VOIGT, C. A., SIMMONS, B. A. & KEASLING, J. D. 2011. Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proceedings of the National Academy of Sciences USA,* 108**,** 19949-19954.
BOLLAG, D. M., ROZYCKI, M. D. & EDELSTEIN, S. J. 1996. *Protein Methods,* New York, Wiley-Liss, Inc.
BORASTON, A. B., BOLAM, D. N., GILBERT, H. J. & DAVIES, G. J. 2004. Carbohydrate-binding modules: fine-tuning polysaccharide recognition. *Biochemical Journal,* 382**,** 769-781.
BORASTON, A. B., BUEREN, A. L. V., FICKO-BLEAN, E. & ABBOTT, D. W. 2007. Carbohydrate–protein interactions: carbohydrate-binding modules. *In:* KAMERLING, H., BOONS, G.-J., LEE, Y. C., SUZUKI, A., TANIGUCHI, N. & VORAGEN, A. G. J. F. (eds.) *Comprehensive Glycoscience: From Chemistry to Systems Biology.* Elsevier B.V.
BOTA, D. A. & DAVIES, K. J. A. 2002. *Lon* protease preferentially degrades oxidized mitochondrial aconitase by an ATP-stimulated mechanism. *Nature Cell Biology,* 4**,** 674-680.
BOUDET, A.-M. 2003. Towards an understanding of the supramolecular organization of the lignified wall. *In:* ROSE, J. K. C. (ed.) *The Plant Cell Wall.* Oxford, UK: Blackwell Publishing Ltd.
BRUNNER, F., WIRTZ, W., ROSE, J. K. C., DARVILL, A. G., GOVERS, F., SCHEEL, D. & NURNBERGER, T. 2002. A ß-glucosidase/xylosidase from the phytopathogenic oomycete, *Phytophthora infestans*. *Phytochemistry,* 59**,** 689-696.
CAMMAROTA, M. C. & SANT'ANNAJR., G. L. 1998. Metabolic blocking of exopolysaccharides synthesis: effects on microbial adhesion and biofilm accumulation. *Biotechnology Letters,* 20**,** 1-4.
CANTAREL, B. L., COUTINHO, P. M., RANCUREL, C., BERNARD, T., LOMBARD, V. & HENRISSAT, B. 2009. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Research,* 37**,** D233-238.
CARVALHO, A. L., GOYAL, A., PRATES, J. A. M., BOLAM_, D. N., GILBERT_, H. J., PIRES, V. N. M. R., FERREIRA, L. S. M. A., PLANAS, A., ROMA˜, M. J. & FONTES, C. M. G. A. 2004. The family 11 carbohydrate-binding module of *Clostridium thermocellum* Lic26A-Cel5E accommodates ß-1,4- and ß-1,3-1,4-mixed linked glucans at a single binding site. *Journal of Biological Chemistry,* 279**,** 34785-34793.
CHAPLIN, M. F. 1986. Monosaccharides. *In:* CHAPLIN, M. F. & KENNEDY, J. F. (eds.) *Carbohydrate Analysis.* Oxford, UK: IRL Press.
CHELIUS, M. K. & TRIPLETT, E. W. 2000. *Dyadobacter fermentans* gen. nov., sp.

nov., a novel Gram-negative bacterium isolated from surface-sterilized *Zea mays* stems. *International Journal of Systematic and Evolutionary Microbiology,* 50**,** 751-758.
CHEN, D. & TEXADA, D. E. 2006. Low-usage codons and rare codons of *Escherichia coli*. *Gene Therapy and Molecular Biology,* 10**,** 1-12.
CHEN, H., LI, X. & LJUNGDAHL, L. G. 1994. Isolation and properties of an extracellular ß-glucosidase from the polycentric rumen fungus *Orpinomyces* sp. Strain PC-2. *Applied and Environmental Microbiology,* 60**,** 64-70.
CHUNG, C. T., NIEMELA, S. L. & MILLER, R. H. 1989. One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proceedings of the National Academy of Sciences USA,* 86**,** 2172-2175.
CHUNG, D.-K., SHIN, D.-H., KIM, B.-W., NAM, J.-K., HAN, I.-S. & NAM, S.-W. 1997. Expression and secretion of *Clostridium thermocellum* endoglucanase A gene (*celA*) in different *Saccharomyces cerevisiae* strains. *Biotechnology Letters,* 19**,** 503-506.
CLARKE, A. J. 1997. *Biodegradation of Cellulose: Enzymology and Biotechnology,* Pennsylvania, U.S., Technomic Publishing Company Inc.
COLLINS, F. M. 1964. Cell wall composition of a marine Cytophaga. *The Australian journal of experimental biology and medical science,* 42**,** 263-265.
COUGHLAN, M. P. 1991. Mechanisms of cellulose degradation by fungi and bacteria. *Animal Feed Science and Technology,* 32**,** 77-100.
CZJZEK, M., BOLAM, D. N., MOSBAH, A., ALLOUCH, J., FONTES, C. M. G. A., FERREIRA, L. M. A., BORNET, O., ZAMBONI, V. R., DARBON, H., SMITH, N. L., BLACK, G. W., HENRISSAT, B. & GILBERT, H. J. 2001. The location of the ligand-binding site of carbohydrate-binding modules that have evolved from a common sequence is not conserved. *Journal of Biological Chemistry,* 276**,** 48580-48587.
DENMAN, S., XUE, G.-P. & PATEL, B. 1996. Characterization of a *Neocallimastix patriciarum* cellulase cDNA (*celA*) homologous to *Trichoderma reesei* cellobiohydrolase II. *Applied and Environmental Microbiology,* 62**,** 1889-1896.
DISCHE, Z. 1962. Color reactions of carbohydrates. *In:* WHISTLER, R. L. & WOLFROM, M. L. (eds.) *Methods in Carbohydrate Chemistry Vol I : Analysis and Preparation of Sugars.* New York, U.S.: Academic Press.
DOI, R. H. 2008. Cellulases of mesophilic microorganisms: cellulosome and noncellulosome producers. *Annals of the New York Academy of Sciences,* 1125**,** 267-279.
DUBOC, P. & MOLLET, B. 2001. Applications of exopolysaccharides in the dairy industry. *International Dairy Journal,* 11**,** 759-768.
ERIKSSON, K.-E. & PETTERSSON, B. 1988. Endo-1,4-ß-glucanases of *Sporotrichum pulverulentum*. *In:* WOOD, W. A. & KELLOGG, S. T. (eds.) *Methods in Enzymology (Vol.160) Biomass Part A: Cellulose and hemicellulose.* London, U.K.: Academic Press Inc.
FAURE, D., DESAIR, J., KEIJERS, V., BEKRI, M. A., PROOST, P., HENRISSAT, B. & VANDERLEYDEN, J. 1999. Growth of *Azospirillum irakense* KBC1 on the

aryl ß-glucoside salicin requires either *salA* or *salB*. *Journal of Bacteriology,* 181**,** 3003-3009.
FAURE, D., HENRISSAT, B., PTACEK, D., BEKRI, M. A. & VANDERLEYDEN, J. 2001. The *celA* gene, encoding a Glycosyl Hydrolase Family 3 ß-glucosidase in *Azospirillum irakense,* is required for optimal growth on cellobiosides. *Applied and Environmental Microbiology,* 67**,** 2380-2383.
FINN, R. D., MISTRY, J., TATE, J., COGGILL, P., HEGER, A., POLLINGTON, J. E., GAVIN, O. L., GUNESEKARAN, P., CERIC, G., FORSLUND, K., HOLM, L., SONNHAMMER, E. L., EDDY, S. R. & BATEMAN, A. 2010. The Pfam protein families database. *Nucleic Acids Research,* 38**,** D211-D222.
FOLLETT, E. A. C. & WEBLEY, D. M. 1965. An electron microscope study of the cell surface of *Cytophaga johnsonii* and some observations on related organisms. *Antonie van Leeuwenhoek,* 31**,** 361-382.
FONTES, C. M. G. A. & GILBERT, H. J. 2010. Cellulosomes: highly efficient nanomachines designed to deconstruct plant cell wall complex carbohydrates. *Annual Review of Biochemistry,* 79**,** 655-681.
FORSBERG, Z., VAAJE-KOLSTAD, G., WESTERENG, B., BUNAES, A. C., STENSTROM, Y., MACKENZIE, A., SØRLIE, M., HORN, S. J. & EIJSINK, V. G. H. 2011. Cleavage of cellulose by a CBM33 protein. *Protein Science,* 20**,** 1479-1483.
FRANCISCO, J. A., STATHOPOULOS, C., WARREN, R. A. J., KILBURN, D. G. & GEORGIOU, G. 1993. Specific adhesion and hydrolysis of cellulose by intact *Escherichia coli* expressing surface anchored cellulase or cellulose binding domains. *Nature Biotechnology,* 11**,** 491-495.
FRENCH, A. D. & JOHNSON, G. P. 2007. Cellulose shapes. *In:* R. MALCOLM BROWN, J. & SAXENA, I. M. (eds.) *Cellulose: Molecular and Structural Biology.* Dordrecht, Netherlands: Springer.
FRENCH, C. E. 2009. Synthetic biology and biomass conversion: a match made in heaven? *Journal of the Royal Society Interface,* 6**,** S547-S558.
FRY, S. C. 1988. *The Growing Plant Cell Wall: Chemical and Metabolic Analysis,* Caldwell, New Jersey, U.S., The Blackburn Press.
FRY, S. C. 2004. Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New Phytologist,* 161**,** 641-675.
FUJII, T., FANG, X., INOUE, H., MURAKAMI, K. & SAWAYAMA, S. 2009. Enzymatic hydrolyzing performance of *Acremonium cellulolyticus* and *Trichoderma reesei* against three lignocellulosic materials. *Biotechnology for Biofuels,* 2**,** 1-8.
FUJINO, T., SUKHUMAVASI, J., SASAKI, T., OHMIYA, K. & SHIMIZU, S. 1989. Purification and properties of an endo-1,4-ß-glucanase from *Clostridium josui*. *Journal of Bacteriology,* 171**,** 4076-4079.
FUJITA, M. & HARADA, H. 2001. Ultrastructure and formation of wood cell wall. *In:* HON, D. N.-S. & SHIRAISHI, N. (eds.) *Wood and Cellulosic Chemistry.* 2nd Ed. ed. New York, U.S.: Marcel Dekker, Inc.
GALAZKA, J. M., TIAN, C., BEESON, W. T., MARTINEZ, B., GLASS, N. L. & CATE, J. H. D. 2010. Cellodextrin transport in yeast for improved biofuel production. *Science,* 330**,** 84-86.

GILAD, R., RABINOVICH, L., YARON, S., BAYER, E. A., LAMED, R., GILBERT, H. J. & SHOHAM, Y. 2003. CelI, a noncellulosomal family 9 enzyme from *Clostridium thermocellum*, is a processive endoglucanase that degrades crystalline cellulose. *Journal of Bacteriology,* 185**,** 391-398.

GILKES, N. R., KILBURN, D. G., R. C. MILLER, J. & WARREN, R. A. J. 1991. Bacterial cellulases. *Bioresource Technology,* 36**,** 21-35.

GILKES, N. R., KILBURN, D. G., ROBERT C. MILLER, J. & WARREN, R. A. J. 1989. Structural and functional analysis of a bacterial cellulase by proteolysis. *Journal of Biological Chemistry,* 264**,** 17802-17808.

GILKES, N. R., WARREN, R. A. J., ROBERT C. MILLER, J. & KILBURN, D. G. 1988. Precise excision of the cellulose binding domains from two *Cellulomonas fimi* cellulases by a homologous protease an the effect on catalysis. *Journal of Biological Chemistry,* 263**,** 10401-10407.

GLAZER, A. N. & NIKAIDO, H. 1998. Biomass. *Microbial Biotechnology: fundamentals of applied microbiology.* 2nd ed. New York, U.S.: W.H. Freeman and Company.

GONZÁLEZ-POMBO, P., PÉREZ, G., CARRAU, F., GUISÁN, J. M., BATISTA-VIERA, F. & BRENA, B. M. 2008. One-step purification and characterization of an intracellular ß-glucosidase from *Metschnikowia pulcherrima*. *Biotechnology Letters,* 30**,** 1469-1475.

GRUNO, M., VALJAMAE, P., PETTERSSON, G. & JOHANSSON, G. 2004. Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate. *Biotechnology and Bioengineering,* 86**,** 503-511.

GUILLÉN, D., SÁNCHEZ, S. & RODRÍGUEZ-SANOJA, R. 2010. Carbohydrate-binding domains: multiplicity of biological roles. *Applied Microbiology and Biotechnology,* 85**,** 1241-1249.

GUSTAFSSON, C., GOVINDARAJAN, S. & MINSHULL, J. 2004. Codon bias and heterologous protein expression. *Trends in Biotechnology,* 22**,** 346-353.

GUTLEBEN, W., UNTERHOLZNER, V., VOLGGER, D. & ZEMANN, A. 2004. Characterization of Carbohydrates in Paper and Paper Pulps Using Anion Exchange Chromatography and Principal Component Analysis. *Microchimica Acta,* 146**,** 111-117.

HAMADA, N., KODAIRA, R., NOGAWA, M., SHINJI, K., ITO, R., AMANO, Y., SHIMOSAKA, M., KANDA, T. & OKAZAKI, M. 2001. Role of cellulose-binding domain of exocellulase I from white rot basidiomycete *Irpex lacteus Journal of Bioscience and Bioengineering,* 91**,** 359-362.

HAN, S. J., YOO, Y. J. & KANG, H. S. 1995. Characterization of a bifunctional cellulase and its structural gene: the cel gene of *Bacillus* sp. D04 has exo- and endoglucanase activity. *Journal of Biological Chemistry,* 270**,** 26012-26019.

HANNIG, G. & MAKRIDES, S. C. 1998. Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology,* 16**,** 54-60.

HARHANGI, H. R., STEENBAKKERS, P. J. M., AKHMANOVA, A., JETTEN, M. S. M., DRIFT, C. V. D. & CAMP, H. J. M. O. D. 2002. A highly expressed family 1 ß-glucosidase with transglycosylation capacity from the anaerobic fungus

Piromyces sp. E2. *Biochimica et Biophysica Acta,* 1574**,** 293-303.
HARRAH, T., PANILAITIS, B. & KAPLAN, D. 2006. Microbial exopolysaccharides. *In:* DWORKIN, M., FALKOW, S., ROSENBERG, E., SCHLEIFER, K.-H. & STACKEBRANDT, E. (eds.) *Prokaryotes: a handbook on the biology of bacteria.* 3rd ed. New York, U.S.A.: Springer Science + Business Media, Inc.
HARRIS, P. J. & STONE, B. A. 2008. Chemistry and molecular organization of plant cell walls. *In:* HIMMEL, M. E. (ed.) *Biomass Recalcitrance.* Oxford, UK: Blackwell Publishing Ltd.
HARRIS, P. V., WELNER, D., MCFARLAND, K. C., RE, E., POULSEN, J.-C. N., BROWN, K., SALBO, R., DING, H., VLASENKO, E., MERINO, S., XU, F., CHERRY, J., LARSEN, S. & LEGGIO, L. L. 2010. Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. *Biochemistry,* 49**,** 3305-3316.
HASELOFF, J. & AJIOKA, J. 2009. Synthetic biology: history, challenges and prospects. *Journal of the Royal Society Interface,* 6**,** S389-S391.
HAYASE, M., MAEKAWA, A., YUBISUI, T. & MINAMI, Y. 2008. Properties, intracellular localization, and stage-specific expression of membrane-bound ß-glucosidase, BglM1, from *Physarum polycephalum*. *The International Journal of Biochemistry & Cell Biology,* 40**,** 2141-2150.
HEINEMANN, M. & PANKE, S. 2006. Synthetic biology—putting engineering into biology. *Bioinformatics, 22***,** 2790-2799.
HENRIKSSON, G. & LENNHOLM, H. 2009. Cellulose and Carbohydrate Chemistry. *In:* EK, M., GELLERSTEDT, G. & HENRIKSSON, G. (eds.) *Pulp and Paper Chemistry and Technology Volume 1: Wood Chemistry and Wood Biotechnology.* Berlin, Germany: Deutsche Nationalbibliothek.
HENRISSAT, B. 1994. Cellulases and their interaction with cellulose. *Cellullose,* 1**,** 169-196.
HENSHAW, J. L., BOLAM, D. N., PIRES, V. N. M. R., CZJZEK, M., HENRISSAT, B., FERREIRA, L. M. A., FONTES, C. M. G. A. & GILBERT, H. J. 2004. The family 6 carbohydrate binding module CmCBM6-2 contains two ligand-binding sites with distinct specificities. *Journal of Biological Chemistry,* 279**,** 21552-21559.
HERNANDEZ-MENA, R. & FRIEND, P. L. 1993. Analysis of microbial exopolysaccharides from industrial water systems. *Journal of Industrial Microbiology,* 12**,** 109-113.
HERVÉ, C., ROGOWSKI, A., BLAKE, A. W., MARCUS, S. E., GILBERT, H. J. & KNOX, J. P. 2010. Carbohydrate-binding modules promote the enzymatic deconstruction of intact plant cell walls by targeting and proximity effects. *Proceedings of the National Academy of Sciences USA,* 107**,** 15293-15298.
HRMOVA, M., MACGREGOR, E. A., BIELY, P., STEWART, R. J. & FINCHER, G. B. 1998. Substrate binding and catalytic mechanism of a barley ß-D-glucosidase/(1,4)-ß-D-glucan exohydrolase. *Journal of Biological Chemistry,* 273**,** 11134-11143.
HUMPHREY, B. A., DICKSON, M. R. & MARSHALL, K. C. 1979. Physicochemical and in situ observations on the adhesion of gliding bacteria to surfaces.

Archives of Microbiology, 120**,** 231-238.
HUTCHINSON, H. B. & CLAYTON, J. 1919. On the decomposition of cellulose by an aerobic organism (*Spirochaeta cytophaga,* N. SP.). *Journal of Agricultural Science,* 9**,** 143-173.
IRWIN, D. C., SPEZIO, M., WALKER, L. P. & WILSON, D. B. 1993. Activity studies of eight purified cellulases: specificity, synergism, and binding domain effects. *Biotechnology and Bioengineering,* 42**,** 1002-1013.
JAHN, T. L. & BOVEE, E. C. 1969. Movement and locomotion of microorganisms. *Anual Review of Microbiology,* 19**,** 21-58.
JIAO, Y., CODY, G. D., HARDING, A. K., WILMES, P., SCHRENK, M., WHEELER, K. E., BANFIELD, J. F. & THELEN, M. P. 2010. Characterization of extracellular polymeric substances from acidophilic microbial biofilms. *Applied and Environmental Microbiology,* 76**,** 2916-2922.
JORK, H., FUNK, W., FISCHER, W. & WIMMER, H. 1990. *Thin-Layer Chromatography: Reagent and Detection Methods, Volume 1a, Physical and Chemical Detection Methods: Fundamentals, Reagents I. ,* Germany, VCH Verlagsgesellschaft.
JORK, H., FUNK, W., FISCHER, W. & WIMMER, H. 1994. *Thin-Layer Chromatography: Reagent and Detection Methods, Volume 1b, Physical and Chemical Detection Methods: Activation Reactions, Reagent Sequence, Reagents II. ,* Germany, VCH Verlagsgesellschaft.
KÄMPFER, P., RAUHOPP, O. & DOTT, W. 1991. Glycosidase profiles of members of the Family Ènterobacteriaceae. *Journal of Clinical Microbiology,* 29**,** 2877-2879.
KANE, J. F. 1995. Effects of rare codon clusters on high-level expression of heterologous proteins in *Escherichia coil*. *Current Opinion in Biotechnology,* 6**,** 494-500.
KARKEHABADI, S., HANSSON, H., KIM, S., PIENS, K., MITCHINSON, C. & SANDGREN, M. 2008. The first structure of a glycoside hydrolase family 61 member, Cel61B from *Hypocrea jecorina*, at 1.6 Å resolution. *Journal of Molecular Biology,* 383**,** 144-154.
KARLSSON, J., SALOHEIMO, M., SIIKA-AHO, M., TENKANEN, M., MERJAPENTTILÄ & TJERNELD, F. 2001. Homologous expression and characterization of Cel61A (EG IV) of *Trichoderma reesei*. *European Journal of Biochemistry,* 268**,** 6498-6507.
KASHIWAGI, Y., IIJIMA, C., SASAKI, T. & TANIGUCHI, H. 1991. Characterization of a ß-glucosidase encoded by a gene from *Cellvibrio gilvus*. *Agricultural Biology and Chemistry,* 55**,** 2553-2559.
KATAEVA, I. A., III, R. D. S., SHAH, A., WEST, L. T., LI, X.-L. & LJUNGDAHL, L. G. 2002. The fibronectin type 3-Like repeat from the *Clostridium thermocellum* cellobiohydrolase CbhA promotes hydrolysis of cellulose by modifying Its surface. *Applied and Environmental Microbiology,* 68**,** 4292-4300.
KEEGSTRA, K. 2010. Plant cell walls. *Plant Physiology,* 154**,** 483-486.
KIM, C.-H. 1995. Characterization and substrate specificity of an endo-ß-1,4-D-glucanase I (Avicelase I) from an extracellular multienzyme complex of *Bacillus circulans*. *Applied and Environmental Microbiology,* 61**,** 959-965.

KIM, S. & DALE, B. E. 2004. Global potential bioethanol production from wasted crops and crop residues. *Biomass and Bioenergy,* 26**,** 361-375.
KIM, Y.-W., KANG, K.-S., KIM, S.-Y. & KIM, I.-S. 2000. Formation of fibrillar multimers of oat ß-glucosidase isoenzymes is mediated by the As-Glu1 monomer. *Journal of Molecular Biology,* 303**,** 831-842.
KING, K. W. & VESSAL, M. I. 1969. Enzymes of the celulase complex. *In:* GOULD, R. F. (ed.) *Cellulase and Their Applications.* Washington, D. C., U.S.: American Chemical Society.
KLEMM, D., HEUBLEIN, B., FINK, H.-P. & BOHN, A. 2005. Cellulose: fascinating biopolymer and sustainable raw material. *Angewandte Chemie International Edition,* 44**,** 3358-3393.
KLEMM, D., PHILIPP, B., HEINZE, T., HEINZE, U. & WAGENKNECHT, W. 1998. General considerations on structure and reactivity of cellulose. *Comprehensive Cellulose Chemistry; Volume 1: Fundamentals and Analytical Methods.* Weinheim, Germany: Wiley-VCH.
KNIGHT, T. 2003. Idempotent Vector Design for Standard Assembly of Biobricks. *MIT Artificial Intelligence Laboratory; MIT Synthetic Biology Working Group*.
KOROPATKIN, N. M. & SMITH, T. J. 2010. SusG: A unique cell-membrane-associated α-amylase from a prominent human gut symbiont targets complex starch molecules. *Structure,* 18**,** 200-215.
KOSEKI, T., MESE, Y., FUSHINOBU, S., MASAKI, K., FUJII, T., ITO, K., SHIONO, Y., MURAYAMA, T. & IEFUJI, H. 2008. Biochemical characterization of a glycoside hydrolase family 61 endoglucanase from *Aspergillus kawachii*. *Applied Microbiology and Biotechnology,* 77**,** 1279-1285.
LAGAERT, S., BELIËN, T. & VOLCKAERT, G. 2009. Plant cell walls: protecting the barrier from degradation by microbial enzymes. *Seminars in Cell & Developmental Biology,* 20**,** 1064-1073.
LAIL, K., SIKORSKI, J., SAUNDERS, E., LAPIDUS, A., RIO, T. G. D., COPELAND, A., TICE, H., CHENG, J.-F., LUCAS, S., NOLAN, M., BRUCE, D., GOODWIN, L., PITLUCK, S., IVANOVA, N., MAVROMATIS, K., OVCHINNIKOVA, G., PATI, A., CHEN, A., PALANIAPPAN, K., LAND, M., HAUSER, L., CHANG, Y.-J., JEFFRIES, C. D., CHAIN, P., BRETTIN, T., DETTER, J. C., SCHÜTZE, A., ROHDE, M., TINDALL, B. J., GÖKER, M., BRISTOW, J., EISEN, J. A., MARKOWITZ, V., HUGENHOLTZ, P., KYRPIDES, N. C., KLENK, H.-P. & CHEN, F. 2010. Complete genome sequence of *Spirosoma linguale* type strain (1T). *Standards in Genomic Sciences,* 2**,** 176-185.
LANGSTON, J. A., SHAGHASI, T., ABBATE, E., XU, F., VLASENKO, E. & SWEENEY, M. D. 2011. Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. *Applied and Environmental Microbiology,* 77**,** 7007-7015.
LARKIN, J. M. 1989. Nonphotosynthetic, nonfruiting gliding bacteria. *In:* STALEY, J. T., BRYANT, M. P., PFENNIG, N. & HOLT, J. G. (eds.) *Bergey's manual of systematic bacteriology.* Baltimore, MD., U.S.: Williams and Wilkins.
LEVY, I., SHANI, Z. & SHOSEYOV, O. 2002. Modification of polysaccharides and plant cell wall by endo-1,4-ß-glucanase and cellulose-binding domains.

Biomolecular Engineering, 19**,** 17-30.
LI, Y., IRWIN, D. C. & WILSON, D. B. 2007. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. *Applied and Environmental Microbiology,* 73**,** 3165-3172.
LI, Y. & WILSON, D. B. 2008. Chitin binding by *Thermobifida fusca* cellulase catalytic domains. *Biotechnology and Bioengineering,* 100**,** 644-652.
LIN, L.-L., RUMBAK, E., ZAPPE, H., THOMSON, J. A. & WOODS, D. R. 1990. Cloning sequencing and analysis of expression of a *Butyrivibrio fibrisolvens* gene encoding a ß-glucosidase. *Journal of General Microbiology,* 136**,** 1567-1576.
LINDER, M. & TEERI, T. T. 1997. The roles and function of cellulose-binding domains. *Journal of Biotechnology,* 57**,** 15-28.
LIU, J. & XIA, W. 2006. Purification and characterization of a bifunctional enzyme with chitosanase and cellulase activity from commercial cellulase. *Biochemical Engineering Journal,* 30**,** 82-87.
LIU, W., BEVAN, D. R. & ZHANG, Y.-H. P. 2009. The Family 1 Glycoside Hydrolase from *Clostridium cellulolyticum* H10 is a cellodextrin glucohydrolase. *Applied Biochemistry and Biotechnology,* 161**,** 264-273.
LIZON, P. & SAMUELS, G. 1997. Taxonomy and nomenclature defined. *Inoculum: Newsletter of the Mycological Society of America.*
LOEWUS, F. A. 1952. Improvement in anthrone method for determination of carbohydrates. *Analytical Chemistry,* 24**,** 219.
LONGONI, P., RODOLFI, M., PANTALEONI, L., DORIA, E., CONCIA, L., PICCO, A. M. & CELLA, R. 2012. Functional analysis of the degradation of cellulosic substrates by a *Chaetomium globosum* endophytic isolate. *Applied and Environmental Microbiology,* 78**,** 3693-3705.
LOUIME, C., ABAZINGE, M. & JOHNSON, E. 2006. Location, formation and biosynthetic regulation of cellulases in the gliding bacteria *Cytophaga hutchinsonii*. *International Journal of Molecular Sciences,* 7**,** 1-11.
LOUIME, C., ABAZINGE, M., JOHNSON, E., LATINWO, L., IKEDIOBI, C. & CLARK, A.-M. 2007. Molecular cloning and biochemical characterization of a family-9 endoglucanase with an unusual structure from the gliding bacteria *Cytophaga hutchinsonii*. *Applied Biochemistry and Biotechnology,* 136**,** 127-138.
LUDWIG, T. G. & GOLDBERG, H. J. V. 1956. The anthrone method for the determination of carbohydrates in foods and in oral rinsing. *Journal of Dental Research,* 35**,** 90-94.
LYND, L. R., WEIMER, P. J., ZYL, W. H. V. & PRETORIUS, I. S. 2002. Microbial cellulose utilization: fundamentals and biotechnology. *Microbiology and Molecular Biology Reviews,* 66**,** 506-577.
MAIDEN, M. C. J., DAVIS, E. O., BALDWIN, S. A., MOORE, D. C. M. & HENDERSON, P. J. F. 1987. Mammalian and bacterial sugar transport proteins are homologous. *Nature,* 325**,** 641-643.
MAKRIDES, S. C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiological Reviews,* 60**,** 512-538.
MANDELS, M. & REESE, E. T. 1959. Induction of cellulase in fungi by cellobiose.

Journal of Bacteriology, 79**,** 816-826.
MARCHLER-BAUER, A., LU, S., ANDERSON, J. B., CHITSAZ, F. & DERBYSHIRE, M. K. 2011. CDD: A conserved domain database for the functional annotation of proteins. *Nucleic Acids Research,* 39**,** 225-229.
MARTIN, H. H., PREUSSER, H.-J. & VERMA, J. P. 1968. Über die oberflächenstruktur von Myxobakterien: II. anionische heteropolysaccharide als baustoffe der schleimhiille von *Cytophaga hutchinsonii* und *Sporocytophaga myxococcoides*. *Archiv fur Mikrobiologie,* 62**,** 72-84.
MAURIZI, M. R. 1992. Proteases and protein degradation in *Escherichia coli*. *Experientia,* 48**,** 178-201.
MCBRIDE, M. J. 2001. Bacterial gliding motility: multiple mechanisms for cell movement. *Annual Review of Microbiology,* 55**,** 49-75.
MCBRIDE, M. J. 2004. *Cytophaga-Flavobacterium* gliding motility. *Journal of Molecular Microbiology and Biotechnology,* 7**,** 63-71.
MCBRIDE, M. J. & BAKER, S. A. 1996. Development of techniques to genetically manipulate members of the Genera *Cytophaga, Flavobacterium, Flexibacter*, and *Sporocytophaga*. *Applied and Environmental Microbiology,* 62**,** 3017-3022.
MCBRIDE, M. J., XIE, G., MARTENS, E. C., LAPIDUS, A., HENRISSAT, B., RHODES, R. G., GOLTSMAN, E., WANG, W., XU, J., HUNNICUTT, D. W., STAROSCIK, A. M., HOOVER, T. R., CHENG, Y.-Q. & STEIN, J. L. 2009. Novel features of the polysaccharide-digesting gliding bacterium *Flavobacterium johnsoniae* as revealed by genome sequence analysis. *Applied and Environmental Microbiology,* 75**,** 6864-6875.
MCKEE, A. E., RUTHERFORD, B. J., CHIVIAN, D. C., BAIDOO, E. K., JUMINAGA, D., KUO, D., BENKE, P. I., DIETRICH, J. A., MA, S. M., ARKIN, A. P., PETZOLD, C. J., ADAMS, P. D., KEASLING, J. D. & CHHABRA, S. R. 2012. Manipulation of the carbon storage regulator system for metabolite remodeling and biofuel production in *Escherichia coli*. *Microbial Cell Factories,* 11**,** 79 [E-pub ahead of print].
MCLEAN, B. W., BORASTON, A. B., BROUWER, D., SANAIE, N., FYFE, C. A., WARREN, R. A. J., KILBURN, D. G. & HAYNES, C. A. 2002. Carbohydrate-binding modules recognize fine substructures of cellulose. *Journal of Biological Chemistry,* 277**,** 50245-50254.
MISKIN, A. & EDBERG, S. C. 1978. Esculin hydrolysis reaction by *Escherichia coli*. *Journal of Clinical Microbiology,* 7**,** 251-254.
MOHAND-OUSSAID, O., PAYOT, S., GUEDON, E., GELHAYE, E., YOUYOU, A. & PETITDEMANGE, H. 1999. The extracellular xylan degradative system in *Clostridium cellulolyticum* cultivated on xylan: evidence for cellfree cellulosome production. *Journal of Bacteriology,* 181**,** 4035-40.
MONTENECOURT, B. S. & EVELEIGH, D. E. 1977. Preparation of mutants of *Trichoderma reesei* with enhanced cellulase production. *Applied and Environmental Microbiology,* 34**,** 777-782.
MOORE, R. L., BASSET, B. B. & SWIFT, M. J. 1979. Developments in the Remazol Brilliant Blue dye-assay for studying the ecology of cellulose decomposition. *Soil Biology & Biochemistry,* 11**,** 311-312.

MORAG, E., BAYER, E. A. & LAMED, R. 1990. Relationship of cellulosomal and noncellulosomal xylanases of *Clostridium thermocellum* to cellulose-degrading enzymes. *Journal of Bacteriology,* 172**,** 6098-6105.
MORANA, A., ESPOSITO, A., MAURELLI, L., RUGGIERO, G., IONATA, E., ROSSI, M. & CARA, F. L. 2008. A novel thermoacidophilic cellulase from *Alicyclobacillus acidocaldarius*. *Protein & Peptide Letters,* 15**,** 1017-1021.
MORRIS, D. L. 1948. Quantitative determination of carbohydrates with Dreywood's anthrone reagent. *Science,* 107**,** 254-255.
MUKHERJI, S. & OUDENAARDEN, A. V. 2009. Synthetic biology: understanding biological design from synthetic circuits. *Nature Reviews,* 10**,** 859-871.
MURBY, M., UHLÉN, M. & STÅHL, S. 1996. Upstream strategies to minimize proteolytic degradation upon recombinant production in *Escherichia coli*. *Protein Expression and Purification,* 7**,** 129-136.
NAKAZAWA, H., OKADA, K., KOBAYASHI, R., KUBOTA, T., ONODERA, T., OCHIAI, N., OMATA, N., OGASAWARA, W., OKADA, H. & MORIKAWA, Y. 2008. Characterization of the catalytic domains of *Trichoderma reesei* endoglucanase I, II, and III, expressed in *Escherichia coli*. *Applied Microbiology and Biotechnology,* 81**,** 681-689.
NAM, E. S., KIM, M. S., LEE, H. B. & AHN, J. K. 2010. β-Glycosidase of *Thermus thermophilus* KNOUC202: gene and biochemical properties of the enzyme expressed in *Escherichia coli*. *Applied Biochemistry and Microbiology,* 46**,** 515-524.
NATAF, Y., YARON, S., STAHL, F., LAMED, R., BAYER, E. A., SCHEPER, T.-H., SONENSHEIN, A. L. & SHOHAM, Y. 2009. Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. *Journal of Bacteriology,* 191**,** 203-209.
NC-IUBMB, Enzyme Nomenclature, Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (http://www.chem.qmul.ac.uk/iubmb/enzyme/).
NEWCOMBE, F. C. 1899. Cellulose-enzymes. *Annals of Botany,* XIII**,** 49-81.
NIELSEN, P. H. & JAHN, A. 1999. Extraction of EPS. *In:* WINGENDER, J., NEU, T. R. & FLEMMING, H.-C. (eds.) *Microbial Extracellular Polymeric Substances.* Berlin, Germany: Springer-Verlag.
NIGAM, A. & AYYAGARI, A. 2007. *Lab Manual in Biochemistry, Immunology and Biotechnology,* New Delhi, India, Tata McGraw-Hill.
NOVAGEN 2006. *KOD hot start DNA polymerase (User protocol TB341 Rev. C 1106)*, EMD Biosciences, Inc, an Alliliate of Merck KGaA, Darmstadt, Germany.
PALA, H., PINTO, R., MOTA, M., DUARTE, A. P. & GAMA, F. M. 2003. Cellulose-binding domains as a tool for paper recycling. *In:* MANSFIELD, S. D. & SADDLER, J. N. (eds.) *Applications of enzymes to lignocellulosics.* WAshington DC, U.S.A.: American Chemical Society.
PALMER, I. & WINGFIELD, P. T. 1995. Preparation and extraction of insoluble (inclusion-body) proteins from *Escherichia coli*. *In:* COLIGAN, C., DUNN, B., PLOEGH, H., SPEICHER, D. & WINGFIELD, P. (eds.) *Current Protocols in Protein Science.* New York, U.S.: John Wiley & Sons, Inc.

PAO, S. S., PAULSON, I. T. & MILTON H. SAIER, J. 1998. Major facilitator superfamily. *Microbiology and Molecular Biology Reviews,* 62**,** 1-34.
PARK, E. Y., ANH, P. N. & OKUDA, N. 2004. Bioconversion of waste office paper to L(+)-lactic acid by the filamentous fungus *Rhizopus oryzae*. *Bioresource Technology,* 93**,** 77-83.
PARK, Y. W. & YUN, H. D. 1999. Cloning of the *Escherichia coli* endo-1-4-D-glucanase gene and identification of its product. *Molecular Genetics and Genomics,* 261**,** 236-241.
PARLIAMENTARY. *RE: Office of Science and Technology (2008).*
PAUL, A. P. A. K. 2008. Microbial extracellular polymeric substances: central elements in heavy metal bioremediation. *Indian Journal of Microbiology,* 48**,** 49-64.
PERRY, J. D., MORRIS, K. A., JAMES, A. L., OLIVER, M. & GOULD, F. K. 2007. Evaluation of novel chromogenic substrates for the detection of bacterial ß-glucosidase. *Journal of Applied Microbiology,* 102**,** 410-415.
PHILIPPIDIS, G. P. 1994. Cellulase production technology. *In:* HIMMEL, M. E., BAKER, J. O. & OVEREND, R. P. (eds.) *Enzymatic conversion of biomass for fuels production.* Washington DC, U.S.A: American Chemical Society.
PIRES, V. N. M. R., HENSHAW, J. L., PRATES, J. A. M., BOLAM, D. N., FERREIRA, L. S. M. A., FONTES, C. M. G. A., HENRISSAT, B., PLANAS, A., GILBERT, H. J. & CZJZEK, M. 2004. The crystal structure of the family 6 carbohydrate binding module from *Cellvibrio mixtus* endoglucanase 5A in complex with oligosaccharides reveals two distinct binding sites with different ligand specificities. *Journal of Biological Chemistry,* 279**,** 21560-21568.
PLUDE, J. L., PARKER, D. L., SCHOMMER, O. J., TIMMERMAN, R. J., HAGSTROM, S. A., JOERS, J. M. & HNASKO, R. 1991. Chemical characterization of polysaccharide from the slime layer of the Cyanobacterium *Microcystis flos-aquae* C3-40. *Applied and Environmental Microbiology,* 57**,** 1696-1700.
QI, M., JUN, H.-S. & FORSBERG, C. W. 2008. *Cel9D*, an atypical 1,4-ß-D-glucan glucohydrolase from *Fibrobacter succinogenes*: characteristics, catalytic residues, and synergistic interactions with other cellulases. *Journal of Bacteriology,* 190**,** 1976-1984.
QUINLAN, R. J., SWEENEY, M. D., LEGGIO, L. L., OTTEN, H., POULSEN, J.-C. N., JOHANSEN, K. S., KROGH, K. B. R. M., JØRGENSEN, C. I., TOVBORG, M., ANTHONSEN, A., TRYFONA, T., WALTER, C. P., DUPREE, P., XU, F., DAVIES, G. J. & WALTON, P. H. 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. *Proceedings of the National Academy of Sciences USA,* 108**,** 15079-15084.
RÄTTÖ, M., SUIHKO, M.-L. & SIIKA-AHO, M. 2005. Polysaccharide-producing bacteria isolated from paper machine slime deposits. *Journal of Industrial Microbiology & Biotechnology,* 32**,** 109-114.
REESE, E. T. 1976. History of the cellulose program at the U.S. Army Natick Development Center. *In:* ELMER L. GADEN, J., MANDELS, M. H., REESE, E. T. & SPANO, L. A. (eds.) *Biotechnology and Bioengineering Symposium:*

Enzymatic Conversion of Cellulosic Materials: Technology and Application. New York: John Wiley & Sons, Inc.
REESE, E. T., SIU, R. G. H. & LEVINSON, H. S. 1950. The biological degradation of soluble cellulose derivatives and its relationship to the mechenism of cellulose hydrolysis. *Journal of Bacteriology,* 59**,** 485-497.
REEVES, A. R., D'ELIA, J. N., FRIAS, J. & SALYERS, A. A. 1996. A *Bacteroides thetaiotaomicron* outer membrane protein that is essential for utilization of maltooligosaccharides and starch. *Journal of Bacteriology,* 178**,** 823-830.
REEVES, A. R., WANG, G.-R. & SALYERS, A. A. 1997. Characterization of four outer membrane proteins that play a role in utilization of starch by *Bacteroides thetaiotaomicron*. *Journal of Bacteriology,* 179**,** 643-649.
REICHENBACH, H. 2006. The order Cytophagales. *In:* DWORKIN, M. (ed.) *The Prokaryotes.* 3rd ed.
RIXON, J. E., FERREIRA, L. M. A., DURRANT, A. J., LAURIE, J. I., HAZLEWOOD, G. P. & GILBERT, H. J. 1992. Characterization of the gene *celD* and its encoded product 1,4-ß-D-glucan glucohydrolase D from *Pseudomonas fluorescens* subsp. *cellulosa*. *Biochemical Journal,* 285**,** 947-955.
ROLLEFSON, J. B., STEPHEN, C. S., TIEN, M. & BOND, D. R. 2011. Identification of an extracellular polysaccharide network essential for cytochrome anchoring and biofilm formation in *Geobacter sulfurreducens*. *Journal of Bacteriology,* 193**,** 1023-1033.
RUBIN, E. M. 2008. Genomics of cellulosic biofuels. *Nature,* 454**,** 841-845.
SAIER.JR., M. H. 2000. Families of proteins forming transmembrane channels. *Journal of Membrane Biology,* 175**,** 165-180.
SAKON, J., IRWIN, D., WILSON, D. B. & KARPLUS, P. A. 1997. Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. *Nature Structural Biology,* 4**,** 810-818.
SAMBROOK, J. & RUSSELL, D. W. 2001. *Molecular Cloning: A Laboratory Manual,* New York, U.S., Cold Spring Harbor Laboratory Press.
SATO, K., NAITO, M., YUKITAKE, H., HIRAKAWA, H., SHOJI, M., MCBRIDE, M. J., RHODES, R. G. & NAKAYAMA, K. 2010. A protein secretion system linked to *Bacteroidete* gliding motility and pathogenesis. *Proceedings of the National Academy of Sciences USA,* 107**,** 276-281.
SATO, K., SAKAI, E., VEITH, P. D., SHOJI, M., KIKUCHI, Y., YUKITAKE, H., OHARA, N., NAITO, M., OKAMOTO, K., REYNOLDS, E. C. & NAKAYAMA, K. 2005. Identification of a new membrane-associated protein that influences transport/maturation of gingipains and adhesins of *Porphyromonas gingivalis*. *Journal of Biological Chemistry,* 280**,** 8668-8677.
SCHÖNERT, S., SEITZ, S., KRAFFT, H., FEUERBAUM, E.-A., ANDERNACH, I., WITZ, G. & DAHL, M. K. 2006. Maltose and maltodextrin utilization by *Bacillus subtilis*. *Journal of Bacteriology,* 188**,** 3911-3922.
SCHWARZ, W. H. 2001. The cellulosome and cellulose degradation by anaerobic bacteria. *Applied Microbiology and Biotechnology,* 56**,** 634-649.
SHEEHAN, J. & HIMMEL, M. 1999. Enzymes, energy, and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol. *Biotechnology Progress,* 15**,** 817-827.

SHETTY, R. P., ENDY, D. & JR, T. F. K. 2008. Engineering BioBrick vectors from BioBrick parts. *Journal of Biological Engineering,* 2**,** 1-12.
SHEWEITA, S. A., ICHI-ISHI, A., PARK, J.-S., LIU, C., LAERCIO M. MALBURG, J. & DOI, R. H. 1996. Characterization of *engF*, a gene for a non-cellulosomal *Clostridium cellulovorans* endoglucanase. *Gene,* 182**,** 163-167.
SHIPMAN, J. A., BERLEMAN, J. E. & SALYERS, A. A. 2000. Characterization of four outer membrane proteins involved in binding starch to the cell surface of *Bacteroides thetaiotaomicron*. *Journal of Bacteriology,* 182**,** 5365-5372.
SHIPMAN, J. A., CHO, K. H., SIEGEL, H. A. & SALYERS, A. A. 1999. Physiological characterization of SusG, an outer membrane protein essential for starch utilization by *Bacteroides thetaiotaomicron*. *Journal of Bacteriology,* 181**,** 7206-7211.
SHOSEYOV, O., SHANI, Z. & LEVY, I. 2006. Carbohydrate binding modules: biochemical properties and novel applications. *Microbiology and Molecular Biology Reviews,* 70**,** 283-295.
SHULAMI, S., RAZ-PASTEUR, A., TABACHNIKOV, O., GILEAD-GROPPER, S., SHNER, I. & SHOHAM, Y. 2011. The L-arabinan utilization system of *Geobacillus stearothermophilus*. *Journal of Bacteriology,* 193**,** 2838-2850.
SINGH, S. M. & PANDA, A. K. 2005. Solubilization and refolding of bacterial inclusion body proteins. *Journal of Bioscience and Bioengineering,* 99**,** 303-310.
SØRENSEN, H. P. & MORTENSEN, K. K. 2005. Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology,* 115**,** 113-128.
STÅHLBERG, J., JOHANSSON, G. & PETTERSSON, G. 1991. A new model for enzymatic hydrolysis of cellulose based on the two-domain structure of cellobiohydrolase I. *Nature Biotechnology,* 9**,** 286-290.
STANIER, R. Y. 1942. The *Cytophaga* group: A contribution to the biology of Myxobacteria. *Microbiology and Molecular Biology Reviews,* 6**,** 143-196.
STEENBAKKERS, P. J. M., HARHANGI, H. R., BOSSCHER, M. W., HOOFT, M. M. C. V. D., KELTJENS, J. T., DRIFT, C. V. D., VOGELS, G. D. & CAMP, H. J. M. O. D. 2003. ß-Glucosidase in cellulosome of the anaerobic fungus *Piromyces* sp. strain E2 is a family 3 glycoside hydrolase. *Biochemical Journal,* 370**,** 963-970.
STONE, B. A. 1958. Biochemistry of cellulose and related polysaccharides. *Nature,* 182**,** 687-690.
STOPPOK, W., RAPP, P. & WAGNER, F. 1982. Formation, location, and regulation of endo-1,4-ß-glucanases and ß-glucosidases from *Cellulomonas uda*. *Applied and Environmental Microbiology,* 44**,** 44-53.
STUBBS, K. A., BALCEWICH, M., MARK, B. L. & VOCADLO, D. J. 2007. Small molecule inhibitors of a glycoside hydrolase attenuate inducible AmpC-mediated beta-lactam resistance. *Journal of Biological Chemistry,* 282**,** 21382-21391.
STURGEON, R. J. 2000. *Monosaccharides and sugar alcohol analysis* [Online]. John Wiley & Sons Ltd.
SUL, O.-J., KIM, J.-H., PARK, S.-J., SON, Y.-J., PARK, B.-R., CHUNG, D. K.,

JEONG, C.-S. & HAN, I.-S. 2004. Characterization and molecular cloning of a novel endoglucanase from *Trichoderma* sp. C-4. *Applied Microbiology and Biotechnology,* 66**,** 63-70.
SUTHERLAND, I. W. 1985. Biothynthesis and composition of Gram-negative bacterial extracellular and wall polysaccharides. *Annual Review of Microbiology,* 39**,** 243-270.
SUTHERLAND, I. W. 1990. Introduction and definition. *Biotechnology of Microbial Exopolysaccharides.* Cambridge, U.K.: Cambridge University Press.
SUTHERLAND, I. W. 1996. Extracellular polysaccharides. *In:* ROEHR, M. (ed.) *Biotechnology: Products of primary metabolism.* 2nd ed. Weinheim, Germany: VCH Verlagsgesellschaft mbH.
SUTHERLAND, I. W. 1997. Microbial exopolysaccharides-structural subtleties and their consequences. *Pure and Applied Chemistry,* 69**,** 1911-1917.
SUTHERLAND, I. W. 1998. Novel and established applications of microbial polysaccharide. *Trends in Biotechnology,* 16**,** 41-46.
SUTHERLAND, I. W. 1999. Biofilm Exopolysaccharides. *In:* WINGENDER, J., NEU, T. R. & FLEMMING, H.-C. (eds.) *Microbial Extracellular Polymeric Substances.* Berlin, Germany: Springer-Verlag.
SUTHERLAND, I. W. 2001. Microbial polysaccharides from Gram-negative bacteria. *International Dairy Journal,* 11**,** 663-674.
SYGMUND, C., KRACHER, D., SCHEIBLBRANDNER, S., ZAHMA, K., FELICE, A. K. G., HARREITHER, W. & KITTL, R. 2012. Characterization of the two *Neurospora crassa* cellobiose dehydrogenases and their connection to oxidative cellulose degradation. *Applied and Environmental Microbiology,* (In press).
TE'O, V. S. J., SAUL, D. J. & BERGQUIST, P. L. 1995. *celA*, another gene coding for a multidomain cellulase from the extreme thermophile *Caldocellum saccharolyticum*. *Applied Microbiology and Biotechnology,* 43**,** 291-296.
TEERI, T. T. 1997. Crystalline cellulose degradation: new insight into the function of cellobiohydrolases. *Trends in Biotechnology,* 15**,** 160-167.
TENKANEN, M., NIKU-PAAVOLA, M.-L., LINDER, M. & VIIKARI, L. 2003. Cellulases in food processing. *In:* WHITAKER, J. R., VORAGEN, A. G. J. & WONG, D. W. S. (eds.) *Handbook of Food Enzymology.* New York, U.S.: Marcel Dekker, Inc.
THOMPSON, J. S., HODGE, D. S. & BORCZYK, A. A. 1990. Rapid biochemical test to identify verocytotoxin-positive strains of *Escherichia coli* serotype O157. *Journal of Clinical Microbiology,* 28**,** 2165-2168.
TOBIAS, J. W., SHRADER, T. E., ROCAP, G. & VARSHAVSKY, A. 1991. The N-End Rule in Bacteria. *Science,* 254**,** 1374-1377.
TOKUDA, G., LO, N., WATANABE, H., SLAYTOR, M., MATSUMOTO, T. & NODA, H. 1999. Metazoan cellulase genes from termites: intron/exon structures and sites of expression. *Biochimica et Biophysica Acta,* 1447**,** 146-159.
TOMME, P., KWAN, E., GILKES, N. R., KILBURN, D. G. & WARREN, R. A. J. 1996. Characterization of CenC, an enzyme from *Cellulomonas fimi* with both endo- and exoglucanase activities. *Journal of Bacteriology,* 178**,** 4216-4223.
TREPETA, R. W. & EDBERG, S. C. 1984. Methylumbelliferyl-ß-D-glucuronide-

based medium for rapid isolation and identification of *Escherichia coli*. *Journal of Clinical Microbiology,* 19**,** 172-174.
TSUNEDA, S., AIKAWA, H., HAYASHI, H., YUASA, A. & HIRATA, A. 2003. Extracellular polymeric substances responsible for bacterial adhesion onto solid surface. *FEMS Microbiology Letters,* 223**,** 287-292.
TURNER, P., PRAMHED, A., KANDERS, E., HEDSTRO¨M, M., KARLSSON, E. N. & LOGAN, D. T. 2007a. Expression, purification, crystallization and preliminary X-ray diffraction analysis of *Thermotoga neapolitana* ß-glucosidase B. *Acta Crystallographica,* F63**,** 802-806.
TURNER, P., SVENSSON, D., ADLERCREUTZ, P. & KARLSSON, E. N. 2007b. A novel variant of *Thermotoga neapolitana* ß-glucosidase B is an efficient catalyst for the synthesis of alkyl glucosides by transglycosylation. *Journal of Biotechnology,* 130**,** 67-74.
VAAJE-KOLSTAD, G., WESTERENG, B., HORN, S. J., LIU, Z., ZHAI, H., SØRLIE, M. & EIJSINK, V. G. H. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. *Science,* 330**,** 219-222.
VARGHESE, J. N., HRMOVA, M. & FINCHER, G. B. 1999. Three-dimensional structure of a barley ß-D-glucan exohydrolase, a family 3 glycosyl hydrolase. *Structure,* 7**,** 179-190.
VARSHAVSKY, A. 1997. The N-end rule pathway of protein degradation. *Genes to Cells,* 2**,** 13-28.
VERMA, J. P. & MARTIN, H. H. 1967. Chemistry and ultrastructure of surface layers in primitive Myxobacteria: *Cytophaga hutchinsonii* and *Sporocytophaga myxococoides*. *Folia Microbiologica,* 12**,** 248-254.
WALKER, E. & WARREN, F. L. 1938. Decomposition of cellulose by *Cytophaga*. I. *Biochemical Journal,* 32**,** 31-43.
WANG, W. K., KRUUS, K. & WU, J. H. D. 1993. Cloning and DNA sequence of the gene coding for *Clostridium thermocellum* cellulase Ss (CelS), a major cellulosome component. *Journal of Bacteriology,* 175**,** 1293-1302.
WARREN, R. A. J. 1996. Microbial hydrolysis of polysaccharides. *Annual Review of Microbiology,* 50**,** 183-212.
WATANABE, T., ITO, Y., YAMADA, T., HASHIMOTO, M., SEKINE, S. & TANAKA, H. 1994. The roles of the C-terminal domain and type III domains of chitinase A1 from *Bacillus circulans* WL-12 in chitin degradation. *Journal of Bacteriology,* 178**,** 4465-4472.
WEI, H., XU, Q., II, L. E. T., BAKER, J. O., TUCKER, M. P. & DING, S.-Y. 2009. Natural paradigms of plant cell wall degradation. *Current Opinion in Biotechnology,* 20**,** 330-338.
WEICKERT, M. J., DOHERTY, D. H., BEST, E. A. & OLINS, P. O. 1996. Optimization of heterologous protein production in *Escherichia coli*. *Current Opinion in Biotechnology,* 7**,** 494-499.
WEINER, R., LANGILLE, S. & QUINTERO, E. 1995. Structure, function and immunochemistry of bacterial exopollysaccharides. *Journal of Industrial Microbiology,* 15**,** 339-346.
WESTERENG, B., ISHIDA, T., VAAJE-KOLSTAD, G., WU, M., EIJSINK, V. G. H., IGARASHI, K., SAMEJIMA, M., STÅHLBERG, J., HORN, S. J. &

SANDGREN, M. 2011. The putative endoglucanase PcGH61D from *Phanerochaete chrysosporium* Is a metal-dependent oxidative enzyme that cleaves cellulose. *PLos One,* 6**,** e27807.
WILSON, D. B. 2008. Three microbial strategies for plant cell wall degradation. *Annals of the New York Academy of Sciences,* 1125**,** 289-297.
WILSON, D. B. 2009. Cellulases. *In:* SCHAECHTER, M. (ed.) *Encyclopedia of Microbiology.* 3rd ed.: Elsevier Inc.
WINGENDER, J., NEU, T. R. & FLEMMING, H.-C. 1999. What are bacterial extracellular polymeric substances? *In:* WINGENDER, J., NEU, T. R. & FLEMMING, H.-C. (eds.) *Microbial Extracellular Polymeric Substances.* Berlin, Germany: Springer-Verlag.
WINOGRADSKY, S. 1929. E'tudes sur la microbiologie du sol. *Annales de l'Institut Pasteur,* 43**,** 549-633.
WRIGHT, R. M., YABLONSKY, M. D., SHALITA, Z. P., GOYAL, A. K. & EVELEIGH, D. E. 1992. Cloning, characterization, and nucleotide sequence of a gene encoding *Microbispora bispora* BglB, a thermostable ß-glucosidase expressed in *Escherichia coli*. *Applied and Environmental Microbiology,* 58**,** 3455-3465.
WULFF-STROBEL, C. R. & WILSON, D. B. 1995. Cloning, sequencing, and characterization of a membrane-associated *Prevotella ruminicola* B14 ß-glucosidase with cellodextrinase and cyanoglycosidase activities. *Journal of Bacteriology,* 177**,** 5884-5890.
XIE, G., BRUCE, D. C., CHALLACOMBE, J. F., CHERTKOV, O., DETTER, J. C., GILNA, P., HAN, C. S., LUCAS, S., MISRA, M., MYERS, G. L., RICHARDSON, P., TAPIA, R., THAYER, N., THOMPSON, L. S., BRETTIN, T. S., HENRISSAT, B., WILSON, D. B. & MCBRIDE, M. J. 2007. Genome sequence of the cellulolytic gliding bacterium *Cytophaga hutchinsonii*. *Applied and Environmental Microbiology,* 73**,** 3536-3546.
XU, Q., ADNEY, W. S., DING, S.-Y. & HIMMEL, M. E. 2007. Cellulases for biomass conversion. *In:* POLAINA, J. & MACCABE, A. P. (eds.) *Industrial Enzymes.* Springer.
XU, Y., JI, X., CHEN, N., LI, P., LIU, W. & LU, X. 2011. Development of replicative *oriC* plasmids and their versatile use in genetic manipulation of *Cytophaga hutchinsonii*. *Applied Microbiology and Biotechnology,* 93**,** 697-705.
YAMANE, K., SUZUKI, H. & MISIZAWA, K. 1970. Purification and properties of extracellular and cell-bound cellulase components of *Pseudomonas fluorescens* var. *cellulosa*. *Journal of Biochemistry,* 67**,** 19-35.
YANG, M., LUOH, S.-M., GODDARD, A., REILLY, D., HENZEL, W. & BASS, S. 1996. The *bglX* gene located at 47.8 min on the *Escherichia coli* chromosome encodes a periplasmic ß-glucosidase. *Microbiology,* 142**,** 1659-1665.
YANG, S., JIANG, Z., YAN, Q. & ZHU, H. 2008. Characterization of a thermostable extracellular ß-glucosidase with activities of exoglucanase and transglycosylation from *Paecilomyces thermophila*. *Journal of Agricultural and Food Chemistry,* 56**,** 602-608.
YU, N. Y., WAGNER, J. R., LAIRD, M. R., MELLI, G., REY, S., LO, R., DAO, P.,

SAHINALP, S. C., ESTER, M., FOSTER, L. J. & BRINKMAN, F. S. L. 2010. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics,* 26**,** 1608-1615.

ZHANG, Y.-H. P., HIMMEL, M. E. & MIELENZ, J. R. 2006. Outlook for cellulase improvement: screening and selection strategies. *Biotechnology Advances,* 24**,** 452-481.

ZHANG, Y.-H. P. & LYND, L. R. 2004. Toward an aggregated understanding of enzymatic hydrolysis of cellulose: noncomplexed cellulase systems. *Biotechnology and Bioengineering,* 88**,** 797-824.

ZHU, Y., LI, H., ZHOU, H., CHEN, G. & LIU, W. 2010. Cellulose and cellodextrin utilization by the cellulolytic bacterium *Cytophaga hutchisonii*. *Bioresource Technology,* 101**,** 6432-6437.

ZVERLOV, V. V., SCHANTZ, N. & SCHWARZ, W. H. 2005. A major new component in the cellulosome of *Clostridium thermocellum* is a processive endo-ß-1,4-glucanase producing cellotetraose. *FEMS Microbiology Letters,* 249**,** 353-358.

Appendix I

Lists of Chemicals and Recipes of Buffers, Reagents and Media

AI-1 Chemical List

Name	Chemical Formula	CAS Number
Acetic acid	$C_2H_4O_2$	9035-69-2
Acetonitrile	C_2H_3N	75-05-8
Acrylamide/bis 30% (Bio-Rad)		
Agarose	$[C_{24}H_{38}O_{19}]_n$	9063-31-4
Ammonium persulfate	$H_8N_2O_8S_2$	7727-54-0
Anthrone	$C_{14}H_{10}O$	90-44-8
Arabinose	$C_5H_{10}O_5$	7296-59-5
Bovine serum albumin		
Bradford reagent (Thermo Scientific)		
Bromophenol blue	$C_{19}H_{10}Br_4O_5S$	632-72-4
Butan-1-ol	$C_4H_{10}O$	82115-62-6
Carboxymethyl Cellulose	$C_8H_{16}NaO_8$	99331-82-5
Cellobiose	$C_{12}H_{22}O_{11}$	9004-34-6
Congo Red	$C_{32}H_{22}N_6Na_2O_6S_2$	87440-95-7
Coomassie Brilliant Blue	$C_{45}H_{44}N_3NaO_7S_2$	86697-48-5
DEAE Sepharose CL6B (Sigma)		
Dimethyl sulfoxide	C_2H_6OS	9008-97-3
Dipotassium hydrogen phosphate	HK_2O_4P	7758-11-4
Ethanol	C_2H_6O	9003-99-0
Ethylenediaminetetraacetic acid	$C_{10}H_{16}N_2O_8$	94108-75-5
Filter paper (Wiggins Teape)		
Fucose	$C_6H_{12}O_5$	87-96-7
Galactose	$C_6H_{12}O_6$	921-60-8
Gel green (Biotium)		
Glucose	$C_6H_{12}O_6$	921-60-8
Glycerol	$C_3H_8O_3$	8043-29-6
Glycine	$C_2H_5NO_2$	87867-94-5

Name	Chemical Formula	CAS Number
Hydrochloric acid	ClH	9004-54-0
Iron(2+) sulfate heptahydrate	$FeH_{14}O_{11}S$	7782-63-0
1-(Isopropylthio)-ß-galactopyranside	$C_9H_{18}O_5S$	367-93-1
Lactose	$C_{12}H_{22}O_{11}$	89466-76-2
Magnesium chloride	Cl_2Mg	7786-30-3
Magnesium sulfate heptahydrate	$H_{14}MgO_{11}S$	7487-88-9
Maltose	$C_{12}H_{22}O_{11}$	9005-84-9
Mannose	$C_6H_{12}O_6$	921-60-8
2-Mercaptoethanol	C_2H_6OS	99748-78-4
Methanol	CH_4O	67-56-1
4-Methylumbelliferone	$C_{10}H_8O_3$	90-33-5
4-Methylumbelliferyl-ß-D-cellobioside	$C_{22}H_{28}O_{13}$	72626-61-0
4-Methylumbelliferyl-ß-D-glucopyranoside	$C_{16}H_{18}O_8$	18997-57-4
Polyethylene Glycol	$[C_2H_6O_2]_n$	71767-64-1
Potassium acetate	$C_2H_3KO_2$	134092-62-9
Potassium chloride	ClK	7447-40-7
Potassium dihydrogen phosphate	H_2KO_4P	7778-77-0
Potassium sodium tartrate (Rochelle salt)	$C_4H_4KNaO_6$	304-59-6
Silica	O_2Si	99439-28-8
Sodium chloride	ClNa	8028-77-1
Sodium dodecyl sulfate (SDS)	$C1_2H_{25}NaO_4S$	8012-56-4
Sodium hydroxide	HNaO	95077-05-7
Sodium iodide	Ina	7681-82-5
Sodium nitrate	$NNaO_3$	7631-99-4
Sodium sulphate	Na_2O_4S	7757-82-6
Sucrose	$C1_2H_{22}O_{11}$	92004-84-7
Sulfuric acid	H_2O_4S	7664-93-9
SybrSafe (Invitrogen)		
N,N,N′,N′-Tetramethylethylenediamine (TEMED)	$C_6H_{16}N_2$	110-18-9
Thymol	$C_{10}H_{14}O$	89-83-8

AI-1 Chemical List (continued)

Name	Chemical Formula	CAS Number
Tris/Base	$C_4H_{11}NO_3$	83147-39-1
Tris/HCl	$C_4H_{12}ClNO_3$	35087-75-3
Tryptone (Fluka)		
Xylose	$C_5H_{10}O_5$	7296-59-5
Yeast extract (Sigma)		

AI-2 Medium and Agar Recipes

Dubos Salt Medium 3 (DSM3)

K_2HPO_4	0.8 g
KH_2PO_4	0.2 g
$MgSO_4 \cdot 7H_2O$	0.5 g
KCl	0.5 g
$NaNO_3$	0.5 g
$FeSO_4 \cdot 7H_2O$	20 mg

pH 7.2 ± 0.2 at 25 °C

Add components to deionized water and bring volume to 1.0 l. Autoclave for 15 min at 15 psi pressure–121 °C. Solid plate is prepared as above with another 1.5 % agar.

Dubos Salt Medium 3 –Tryptone (DSM3T)

All chemicals are prepared as DSM3 except 1.0 g tryptone is added in 1.0 l medium.

Luria-Bertani (LB) Medium

NaCl	10.0 g
Pancreatic digest of casein	10.0 g
Yeast extract	5.0 g

pH 7.0 ± 0.2 at 25 °C

Add components to deionized water and bring volume to 1.0 l. Autoclave for 15 min at 15 psi pressure–121 °C. Solid plate is prepared as above with another 1.5 % agar.

PBS Agar

M9 Medium

K_2HPO_4	0.8 g
KH_2PO_4	0.2 g
$MgSO_4 \cdot 7H_2O$	0.5 g
KCl	0.5 g
$NaNO_3$	0.5 g
$FeSO_4 \cdot 7H_2O$	20 mg

pH 7.2 ± 0.2 at 25 °C

Add components to deionized water and bring volume to 1.0 l. Autoclave for 15 min at 15 psi pressure–121 °C. Solid plate is prepared as above with another 1.5 % agar.

PBS (pH 7.0±0.2)	1.0 l

pH 7.0 ± 0.2 at 25 °C

Add 1.5 % agar. Autoclave for 15 min at 15 psi pressure–121 °C.
Substrate supplement: Add 2.0 g of CMC and autoclave directly.

AI-3 Buffer and Reagent

Phosphate Buffer Saline (1xPBS)

KH_2PO_4	17.0 ml
NaCl	5.0 ml
Na_2HPO_4	1.0 ml

pH 7.0 ± 0.2 at 25 °C

Dissolved and bring to 1.0 l water. Autoclave for 15 min at 15 psi pressure–121 °C.

SDS-PAGE Running Buffer (5x)

Tris base	9.0 g
Glycine	43.2 g
SDS	3.0 g

pH 8.6

Dissolve in deionized water and bring final volume to 600 ml

SDS-PAGE De-Stain

Methanol	40 %
Acetic acid	10 %

Gel is de-stained for overnight.

SDS-PAGE Sample Reagent

Deionized water	3.8 ml
0.5 M Tris-HCl, pH 6.8	1.0 ml
Glycerol	0.8 ml
10% (w/v) SDS	1.6 ml
2-mercaptoethanol	0.4 ml
1% (w/v) bromophenol blue	0.4 ml

Dilute the sample at least 1:4 with sample buffer, and heat at 95 °C for 4 minutes.

SDS-PAGE Stain

Coomassie brilliant blue	0.1 %
Methanol	40 %
Acetic acid	10 %

Gel is stained at least for 30 min.

Transformation and Storage Solution (TSS)

LB medium	17.0 ml
PEG 3350 [40 % (w/v)]	5.0 ml
$MgCl_2$ [1M]	1.0 ml
DMSO	1.0 ml

All components apart from DMSO are autoclaved prior to mixing. TSS is stored at 4 ºC.

Solution 1, 2 and 3

Solution 1	5 µl/ml of 5 mg/ml RNAse A in water
Solution 2	For 6 miniprep: 0.56 ml of steriled water 0.7 ml of 0.4 M NaOH 0.14 ml of 10 % SDS
Solution 3	29.45 g potassium acetate 11.45 ml acetic acid Mix with steriled water and bring final volume to 100 ml.

AI-4 Antibiotics Stock Concentration

Name	Concentration in Medium	Stock Concentartion
Ampicillin	80 µg/ml	100 mg/ml in H_2O
Carbenicillin	80 µg/ml	100 mg/ml in H_2O
Chloramphenicol	34 µg/ml	50 mg/ml in ethanol

Appendix II Lists of primers

AII-1 Primers for cellulase cloning

Name	Sequence (5'→ 3')
chu1107f	ATGAATATACGTTCGACTCTGC
chu1107r	TTATTATTGTTTAATAACAGACTTTGT
chu1280f	ATGGCAGCATTGGCTTTCG
chu1280r	TTATTACTTTCTGCCTTTTGTTGCC
chu1655f	ATGACAAAAAATATGAAAAAAATAG
chu1655r	TTATTACTTAATTACAGTGAATTTCTGAAC
chu1727f	ATGAAAATACTACACCTGATACC
chu1727r	TTATTATTTCTTTTTCAAATGCAGAGCTTCAGCCATTCC
chu1842f	ATGACTATCAGAACAATCTTTATTTTAATTG
chu1842r	TTATTAAGGGAGTTTTGATTTGAATAA
chu2103f	ATGATTAAAAAAATATCCGTAGTG
chu2103r	TTATTATTTTGTTTTTGATTTCTTTTTAG
chu2149f	ATGAAAAAAAGTTTACTGCTCTTTG
chu2149r	TTATTATTTTTTTATTATTTTCCGGG
chu2235f	ATGAAGCAATTAATAAGTGTATTCC
chu2235r	TTATTATTTTGCGTAAAATTTATTG
chu2268f	ATGAAAAAAATAACCGTATTGATTTCCATCTGGCT
chu2268r	GCTACTAGTATTATTACTCATTAAAATATATTTCTGTC
chu2273f	ATGAAAAAAACAATTTATCTTTTTGG
chu2273r	TTATTACTTTTCTTTGTAGTAAAAATTCAATG
chu3577f	ATGAATAAAAAACAGCTCCTC
chu3577r	TTATTATTTATAATAAATAGTTGTTTTTTG
chu3784f	ATGAAGCATCTGTATTTTTTCAC
chu3784r	TTATTATTTTATTTCAAAGGATTGTTTC
chu3811f	ATGATACTTTCTGAGCCTAAAGC
chu3811r	TTATTATTCCCGAATAAAAGAACCATACCAGAGACC

AII-2 Primers for mutagenesis

Name	Sequence (5'→ 3')
chu1107m1m2f	TATAGCGGTACAACGTTATTGAACTCAGAC
chu1107m1m2r	AAATTCCACTTTTGAGATTGTACCGTTTGCTAC
chu1107m3f	ACATTAACCGGTACAGGTTCTGCCGTGGG
chu1107m3r	AACCGATGTTTCTGTAGAAATAGCTATTGTATTGC
chu1280m1f	CAGTGGGCTGCAATCGAATTATTTATTGCAACA
chu1280m1r	GAACTCATCTGTAACGTTTCCATCATCATATGC
chu1280m2f	GCAACAGCCACAAGACCGGAAGGAACAAAA
chu1280m2r	CGCAGAAGCATGAACAATTACCTGATCGTC
chu1280m3f	GCTAACCAGGGTATGTTGTTGATGAATGCATAT
chu1280m3r	CGCAGTTGAGTTACTTCCCCAGTTAAAATTACC

chu1280m4f	GCAATTTCAAACTTAGATTATTTATTGGGCCGT
chu1280m4r	CGCAGTCAGGTAAGAACCTTCATTGCTTAATTT
chu1655m1f	CCAAATTCAGGTCAGCAGGATAATTGCC
chu1655m1r	CCCGCCTACCAATAAGCCCGGTACA
chu1842m1f	TTCACTCATCAGGGGGCAAGCTGGAT
chu1842m1r	CTCCATCGGCAAATAAAAATGAAATTCATATGCAA
chu1842m2f	TATAAATCTGCGGCGTGGGACTGGATCGATCA
chu1842m2r	TACATACGGAGTAGCATCATCTTCAAAAATCTGAT
chu1842m3f	GCTGACGCTGCTACAGGGCATTCAAATTCAAACTC
chu1842m3r	GCCCGCAGTACCAGATCCGTTTTCACTCCATGAAT
chu2103m1f	CGCAACGGTTCTTTTCTGAGATGGCA
chu2103m1r	CCGCAGCAGGATTCTTCTG
chu2103m2f	AATAAGGACATAAAGCCGTATGCAGAAGGTGT
chu2103m2r	CCAGGAAGTTGCTTGCAGTGGTTCATTATAAAC
chu2149m1f	CTACAGCGCTTTGTTGCCATGTGTGCTTCA
chu2149m1r	TACGGCATACGAACATTTCATATTGTTTGCATCTT
chu2149m2f	GCGGGCTATGTATCGGTAG
chu2149m2r	AGGCTGCGTGCATGCTGTCT
chu2235m1f	GGGTGCTATTGCTGAAGGACCAGGCGATAAAAAAA
chu2235m1r	GCAGCAGATAATTCGGGATGCAGTGTATACATTTG
chu3811m1f	CCCTTGCAGAATATTCATCAGTATATT
chu3811m1r	TTTTAAGGTATCAATCGGATAGCCCATAC

AII-3 Primers for pT7-7 construction

Name	Sequence (5'→ 3')
chu2268T7-7f	GGCATTCCATATGACAGAAGCGGGGAGTAAACCTC
chu2273T7-7f	GTTTGAATTCCAGCAGCGCCTGAAACGTCA
chu3577T7-7f	GAAAAAACATATGCCAGCAGCAGATGCCATTGAAA
chu3784T7-7f	GCCGGAATTCTCAGCATCGAAAATAAAATTCAT
chu3577T7-7r	CGGCGAATTCTTATTATTTATAATAAATAGTTGTT

Appendix III

Supplements of protocols

AIII-1 DNA purification

The DNA sample from PCR was purified using silica glass beads, which was prepared in 5 ml of sodium iodide (3 M). The DNA sample was added with 3 volumes of 6 M sodium iodide and appropriate volumes of well-mixed glass beads. This mixture was incubated on an ice bath for at least 10 min and then centrifuged at 11000 *g* for 2 min. The supernatant was removed and the glass beads were washed with ice-cold wash buffer, which contains 10 mM Tris/HCl (pH 7.5), 2.5 mM EDTA, 50 mM sodium chloride and 50 % (v/v) ethanol. The sample was then centrifuged briefly at 11000 *g* and the supernatant was removed. The pellet was re-suspended with appropriate volumes of elution buffer (10 mM Tris, pH 8.0) and the mixture was incubated in a 55 ºC water bath for 10 min. The mixture was then centrifuged at 11000 *g* for 2 min. The purified DNA was collected from the supernatant.

The QIAGEN kit was used to purify the DNA from the agarose gel and the protocols were applied using QIAGEN's manual.

AIII-2 T4 DNA ligation

The cohesive end DNA ligation was performed by mixing 10 % (v/v) T4 ligase (Promega), 10 % (v/v) T4 ligase buffer (Promega) and 80 % (v/v) DNA sample. As for the blunt end ligation, 5 % (v/v) T4 ligase and 5 % (v/v) T4 polynucleotide kinase were used. The ratios of T4 ligase buffer and DNA sample were the same as

the cohesive end ligation. Both ligations were incubated overnight at 16 ºC. After ligation, the ligase was de-activated at 65 ºC for 20 min.

AIII-3 Transformation

Competent cells preparation

Five bacteria strains were used in this research: *Escherichia coli* JM109, *E. coli* JW2120−1, *E. coli* Rosetta, *E. coli* BL21(DE3)pLysS and *Citrobacter freundii*. The competent cell preparation of all these strains was followed the protocols of Chung et al. (1989). A 5 ml starter culture was incubated overnight and 1 ml of this was transferred to a new 50 ml LB medium with appropriate antibiotics. This culture was incubated at 37 ºC until the OD_{600} reached 0.5 or so. The culture was then chilled in an ice bath and centrifuged at 2000 *g* for 20 min. The supernatant was discarded and the pellet was re-suspended with 5ml of iced TSS (Appendix AI-3) at 4 ºC. The mixture was chilled in an ice bath for 30 min. The mixture was finally transferred into pre-chilled 1.5 ml centrifuge tubes (200 µl mixture per tube) and preserved in -80 ºC freezer.

DNA Transformation procedures

The transformation procedure was based on the report of Chung et al. (1989). Plasmid DNA (1 µl) or ligation product (5 µl) was added into 100 µl ice-cold competent cells. After 40 min incubation in an ice bath, the tube was heat-shocked at 42 ºC for 1 min exactly and then put back in the ice bath quickly for another 2 min incubation. The tube was mixed with 900 µl LB and incubated at 37 ºC for 1 h. The

cells were then spun down, re-suspended in 100 µl LB and spread evenly on an LB plate for overnight incubation.

AIII-4 Plasmid DNA miniprep

The plasmid DNA miniprep procedures were followed the book by Sambrook and Russell (2001). This procedure was applied only with *E. coli* JM109.

A transformant was incubated at 37 ºC for 12 to 16 h in a 5 ml LB culture. A maximum volume of 1.3 ml from this culture was transferred into a 1.5 ml microcentrifuge tube and centrifuged at 4000 *g* for 3 min. After discarding the supernatant, the pellet was re-suspended in 100 µl Solution 1 (Appendix I). Then, 200 µl Solution 2 was added (Appendix I) into the tube with gently flicks. The tube was finally added with 150 µl Solution 3 (Appendix I) and mixed with several times of rotation. After centrifugation at 11000 *g* for 10 min, up to 420 µl of the supernatant was carefully transferred to a new 1.5 ml microcentrifuge tube without any pellet or floating debris. The supernatant was mixed with 920 µl of pure ethanol and incubated at least 10mins in an ice bath. The tube was then centrifuged at 11000 *g* for 10 min and the supernatant was carefully removed without touching the pellet. The pellet was washed with 200 µl of 70 % (v/v) ethanol and cleared by removing the ethanol with brief centrifugation. The plasmid DNA was finally prepared by dissolving the ethanol-free pellet in 40 µl elution buffer.

Appendix IV

GH families related to cellulases (CAZy)

Glycoside Hydrolase Family 1	
Known Activities	**β-glucosidase (EC 3.2.1.21)**; β-galactosidase (EC 3.2.1.23); β-mannosidase (EC 3.2.1.25); β-glucuronidase (EC 3.2.1.31); β-D-fucosidase (EC 3.2.1.38); phlorizin hydrolase (EC 3.2.1.62); **exo-β-1,4-glucanase (EC 3.2.1.74)**; 6-phospho-β-galactosidase (EC 3.2.1.85); 6-phospho-β-glucosidase (EC 3.2.1.86); strictosidine β-glucosidase (EC 3.2.1.105); lactase (EC 3.2.1.108); amygdalin β-glucosidase (EC 3.2.1.117); prunasin β-glucosidase (EC 3.2.1.118); raucaffricine β-glucosidase (EC 3.2.1.125); thioglucosidase (EC 3.2.1.147); β-primeverosidase (EC 3.2.1.149); isoflavonoid 7-O-β-apiosyl-β-glucosidase (EC 3.2.1.161); hydroxyisourate hydrolase (EC 3.-.-.-); β-glycosidase (EC 3.2.1.-)
Mechanism	Retaining
Clan	GH-A
3D Structure Status	(β / α) 8
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental); absent in plant myrosinases

Glycoside Hydrolase Family 3	
Known Activities	**β-glucosidase (EC 3.2.1.21)**; xylan 1,4-β-xylosidase (EC 3.2.1.37); β-N-acetylhexosaminidase (EC 3.2.1.52); glucan 1,3-β-glucosidase (EC 3.2.1.58); **glucan 1,4-β-glucosidase (EC 3.2.1.74)**; exo-1,3-1,4-glucanase (EC 3.2.1.-); alphα-L-arabinofuranosidase (EC 3.2.1.55).
Mechanism	Retaining
Catalytic Nucleophile/Base	Asp (experimental)
Catalytic Proton Donor	Glu (experimental)

Glycoside Hydrolase Family 5	
Known Activities	chitosanase (EC 3.2.1.132); β-mannosidase (EC 3.2.1.25); **Cellulase (EC 3.2.1.4)**; glucan 1,3-β-glucosidase (EC 3.2.1.58); licheninase (EC 3.2.1.73); glucan endo-1,6-β-glucosidase (EC 3.2.1.75); mannan endo-β-1,4-mannosidase (EC 3.2.1.78); endo-β-1,4-xylanase (EC 3.2.1.8); **cellulose β-1,4-cellobiosidase (EC 3.2.1.91)**; endo-β-1,6-galactanase (EC 3.2.1.-); β-1,3-mannanase (EC 3.2.1.-); xyloglucan-specific endo-β-1,4-glucanase (EC 3.2.1.151); mannan transglycosylase (EC 2.4.1.-)
Mechanism	Retaining
Clan	GH-A
3D Structure Status	(β / α) 8
Catalytic Nucleophile/Base	Glu (experimental)

Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family A.

Glycoside Hydrolase Family 6	
Known Activities	**endoglucanase (EC 3.2.1.4); cellobiohydrolase (EC 3.2.1.91)**
Mechanism	Inverting
Catalytic Nucleophile/Base	Asp (experimental)
Catalytic Proton Donor	Asp (experimental)
Note	Formerly known as cellulase family B. The cellobiohydrolases of this family are widely believed to act processively from the non-reducing ends of cellulose chains to generate cellobiose.

Glycoside Hydrolase Family 7	
Known Activities	**endo-β-1,4-glucanase (EC 3.2.1.4); reducing end-acting cellobiohydrolase (EC 3.2.1.176)**; chitosanase (EC 3.2.1.132); endo-β-1,3-1,4-glucanase (EC 3.2.1.73)
Mechanism	Retaining
Clan	GH-B
3D Structure Status	β-jelly roll
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family C. The cellobiohydrolases of this family act processively from the reducing ends of cellulose chains to generate cellobiose. This is markedly different from the IUBMB definition of cellobiohydrolases (EC 3.2.1.91), which act from the non-reducing ends of cellulose.

Glycoside Hydrolase Family 8	
Known Activities	chitosanase (EC 3.2.1.132); **cellulase (EC 3.2.1.4)**; licheninase (EC 3.2.1.73); endo-1,4-β-xylanase (EC 3.2.1.8); reducing-end-xylose releasing exo-oligoxylanase (EC 3.2.1.156)
Mechanism	Inverting
Clan	GH-M
3D Structure Status	(α / α) 6
Catalytic Nucleophile/Base	Asp (inferred)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family D

Glycoside Hydrolase Family 9	
Known Activities	**endoglucanase (EC 3.2.1.4); cellobiohydrolase (EC 3.2.1.91); β-glucosidase (EC 3.2.1.21)**
Mechanism	Inverting
3D Structure Status	(α / α) 6
Catalytic Nucleophile/Base	Asp (experimental)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family E.

Glycoside Hydrolase Family 10	
Known Activities	endo-1,4-b-xylanase (EC 3.2.1.8); endo-1,3-b-xylanase (EC 3.2.1.32)

Mechanism	Retaining
Clan	GH-A
3D Structure Status	(b/a)8
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family F. **Five EC 3.2.1.4 enzymes** are found.

Glycoside Hydrolase Family 12	
Known Activities	**endoglucanase (EC 3.2.1.4)**; xyloglucan hydrolase (EC 3.2.1.151); β-1,3-1,4-glucanase (EC 3.2.1.73); xyloglucan endotransglycosylase (EC 2.4.1.207)
Mechanism	Retaining
Clan	GH-C
3D Structure Status	β-jelly roll
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family H.

Glycoside Hydrolase Family 18	
Known Activities	chitinase (EC 3.2.1.14); endo-b-N-acetylglucosaminidase (EC 3.2.1.96); xylanase inhibitor; concanavalin B; narbonin
Mechanism	Retaining
Clan	GH-K
3D Structure Status	(b/a)8
Catalytic Nucleophile/Base	carbonyl oxygen of C-2 acetamido group of substrate
Catalytic Proton Donor	Glu (experimental)
Note	Contains chitinases of classes III and V. Contains non-catalytic proteins such as xylanase inhibitor; concanavalin B; narbonin. **Four EC3.2.1.4 enzymes** are found

Glycoside Hydrolase Family 19	
Known Activities	chitinase (EC 3.2.1.14); lysozyme (EC 3.2.1.17)
Mechanism	Inverting
Clan	none
Catalytic Nucleophile/Base	Not known
Catalytic Proton Donor	Not known
Note	Contains chitinases of classes I, II, and IV. **One EC 3.2.1.4 enzyme** is found

Glycoside Hydrolase Family 26	
Known Activities	b-mannanase (EC 3.2.1.78); b-1,3-xylanase (EC 3.2.1.32)
Mechanism	Retaining
Clan	GH-A
3D Structure Status	(b/a)8
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental)
Note	Formerly known as cellulase family I. **Four EC 3.2.1.4 enzymes** are found.

Glycoside Hydrolase Family 30	
Known Activities	glucosylceramidase (EC 3.2.1.45); β-1,6-glucanase (EC

	3.2.1.75); β-xylosidase (EC 3.2.1.37); **β-glucosidase (3.2.1.21)**
Mechanism	Retaining
Clan	GH-A
3D Structure Status	(β / α) 8
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (inferred)

Glycoside Hydrolase Family 44	
Known Activities	**endoglucanase (EC 3.2.1.4)**; xyloglucanase (EC 3.2.1.151)
Mechanism	Retaining
3D Structure Status	(β / α) 8
Catalytic Nucleophile/Base	Glu
Catalytic Proton Donor	Glu
Note	Formerly known as cellulase family J. Note that the stereochemistry of the reaction has recently been corrected : Kitago et al., J. Biol. Chem. 282(2007)35703-11 (PMID: 17905739)

Glycoside Hydrolase Family 45	
Known Activities	**endoglucanase (EC 3.2.1.4)**
Mechanism	Inverting
Catalytic Nucleophile/Base	Asp (experimental)
Catalytic Proton Donor	Asp (experimental).
Note	Formerly known as cellulase family K; distantly related to plant expansins

Glycoside Hydrolase Family 48	
Known Activities	**reducing end-acting cellobiohydrolase (EC 3.2.1.176); endo-b-1,4-glucanase (EC 3.2.1.4);** chitinase (EC 3.2.1.14)
Mechanism	Inverting
Clan	GH-M
3D Structure Status	(α / α) 6
Catalytic Nucleophile/Base	Not known
Catalytic Proton Donor	Glu
Note	Formerly known as cellulase family L. Some cellobiohydrolases of this family have been reported to act from the reducing ends of cellulose (EC 3.2.1.-), while others have been reported to operate from the non-reducing ends to liberate cellobiose or cellotriose or cellotetraose (EC 3.2.1.-). This family also contains endo-processive cellulases (EC 3.2.1.-), whose activity is hard to distinguish from that of cellobiohydrolases.

Glycoside Hydrolase Family 51	
Known Activities	α-L-arabinofuranosidase (EC 3.2.1.55); **endoglucanase (EC 3.2.1.4)**
Mechanism	Retaining
Clan	GH-A
3D Structure Status	(β / α) 8 (inferred)

Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Glu (experimental)

Glycoside Hydrolase Family 61	
Known Activities	**endoglucanase (EC 3.2.1.4) [the activity is recently corrected to copper-dependent polysaccharide monooxygenases]**
Mechanism	Not known [recently changed to monooxygenase]
Catalytic Nucleophile/Base	Not known
Catalytic Proton Donor	Not known
Note	The enzymes in this family were originally classified as a glycoside hydrolase family based on measurement of very weak endo-1,4-b-D-glucanase activity in one family member. The structure and mode of action of these enzymes are certainly non-canonical and they cannot be considered as bona fide glycosidases. However, they are kept in the CAZy classification on the basis of their capacity to enhance the breakdown of lignocellulose when used in conjunction with a cellulase or a mixture of cellulases.

Glycoside Hydrolase Family 74	
Known Activities	**endoglucanase (EC 3.2.1.4)**; oligoxyloglucan reducing end-specific cellobiohydrolase (EC 3.2.1.150); xyloglucanase (EC 3.2.1.151)
Mechanism	Inverting
3D Structure Status	7-fold β-propeller
Catalytic Nucleophile/Base	Asp (experimental)
Catalytic Proton Donor	Asp (experimental)

Glycoside Hydrolase Family 116	
Known Activities	acid β-glucosidase (EC 3.2.1.45); **β-glucosidase (EC 3.2.1.21)**; β-xylosidase (EC 3.2.1.37)
Mechanism	Retaining
3D Structure Status	
Catalytic Nucleophile/Base	Glu (experimental)
Catalytic Proton Donor	Asp (experimental)

Glycoside Hydrolase Family 124	
Known Activities	**endoglucanase (EC 3.2.1.4)**
Mechanism	inverting
Note	Created after Bras et al. (2011) Proc.Natl.Acad.Sci.USA; distantly related to lytic transglycosylases of family GH23